CrossMark

Robust surface normal estimation via greedy sparse regression

Mingjing Zhang and Mark S. Drew*

Abstract

Photometric stereo (PST) is a widely used technique of estimating surface normals from an image set. However, it often produces inaccurate results for non-Lambertian surface reflectance. In this study, PST is reformulated as a sparse recovery problem where non-Lambertian errors are explicitly identified and corrected. We show that such a problem can be accurately solved via a greedy algorithm called orthogonal matching pursuit (OMP). The performance of OMP is evaluated on synthesized and real-world datasets: we found that the greedy algorithm is overall more robust to non-Lambertian errors than other state-of-the-art sparse approaches with little loss of efficiency. Along with providing an overview of current methods, novel contributions in this paper are as follows: we propose an alternative sparse formulation for PST; in previous PST studies (Wu et al., Robust photometric stereo via low-rank matrix completion and recovery, 2010), (S. Ikehata et al., Robust photometric stereo using sparse regression, 2012), the surface normal vector and the error vector are treated as two entities and are solved independently. In this study, we convert their formulation into a new canonical form of the sparse recovery problem by combining the two vectors into one large vector in a new "stacked" formulation in this domain. This allows for a large repertoire of existing sparse recovery algorithms to be more straightforwardly applied to the PST problem. In our application of the OMP greedy algorithm, we show that greedy solvers can indeed be applied, with this study supplying the first of such attempt at employing greedy approaches to estimate surface normals within the framework of PST. We numerically compare the performance of several normal vector recovery methods. Most notably, this is the first detailed test on complex images of the normal estimation accuracy of our previously proposed method, least median of squares (LMS).

Keywords: Photometric stereo, Robust, Surface normals, OMP

1 Introduction

Shading in 2D images provides a valuable visual cue for understanding the spatial structure of objects. Photometric Stereo (PST) is a powerful technique that exploits shading information to directly estimate the 3D surface orientation, i.e. normal vectors. In the classical PST problem, the input is a set of n images captured from a fixed viewpoint under n different calibrated lighting conditions; hence, there are n observations of luminance at each pixel location. Under the assumption of a Lambertian reflectance model, where the observed luminance is proportional to the cosine of the incident angle and remains constant regardless of the viewing angle, the relationship

between n observations $\mathbf{y} \in \mathbb{R}^n$ at each pixel and the collection of n lighting directions $\mathbf{L} \in \mathbb{R}^{n \times 3}$ is formulated as a linear equation group with respect to the normal vector $\mathbf{n} \in \mathbb{R}^3$, i.e.,

$$\mathbf{y} = \mathbf{Ln}. \tag{1}$$

We emphasize that there are indeed such a set of n equation at each pixel. In PST, the linear system Eq. 1 is solved via ordinary least squares (LS). The advantage of PST over 3D laser scanning is that the former provides a very high resolution (depending on the actual resolution of the camera) and therefore can capture the fine details of the surface that may not show up in the scanned model. In addition, PST only requires a simple and inexpensive hardware setup whereas 3D scanning devices are usually costly and less portable. Innovative recent work [3, 4] can reduce PST to a single-shot scenario in a different setup

*Correspondence: mark@cs.sfu.ca
School of Computing Science, Simon Fraser University, Vancouver, BC V5A 1S6, Canada

Springer

with spectral multiplexing and more than three colour channels or polarized illumination.

Although the classical PST method almost always guarantees a visually plausible normal map, it in fact suffers from a serious accuracy problem: the simple Lambertian reflectance model adopted in PST does not strictly apply to most real-world textures, which exhibit specular reflection properties to various degrees. Even if the surface is indeed approximately Lambertian, other non-Lambertian errors can be introduced by the interaction of the light and the objects' geometry, resulting in cast shadows, including self-shadowing, as well as interreflections. Attached shadows are also outside the simple shading model. Such non-Lambertian observations, regarded as "outliers" in a Lambertian-based linear model, may severely reduce the accuracy of LS results. Hence, a PST method that is robust to such non-Lambertian effects is needed in order to generate a high-quality normal map.

Many improved PST methods have been proposed since the original PST in an attempt to minimize the effect of non-Lambertian components. These methods either adopt a more sophisticated reflectance model to accommodate non-Lambertian observations as "inliers" (e.g. [5–7]) or rather keep the Lambertian model but use robust statistical methods to rule out or reduce the effect of non-Lambertian outliers (e.g. [8–10]). A typical example of the second category is the least median of squares (LMS) approach used in our previous study

[11] (and see [10, 12]), in which the observations outside a certain confidence band are deemed to be outliers. In this study, we again adopt the Lambertian model, but solve for the normal vectors via a sparse representation framework that estimates both the normals and non-Lambertian errors at the same time. This sparse method is more closely related to the statistical-based methods.

1.1 Sparse representation and recovery

It is well understood that ordinary LS fails to unambiguously reconstruct a signal that is passed through an underdetermined linear system, where the number of unknown variables exceeds that of linear equations (Fig. 1b). However, it has been shown that if the signal to be recovered is sparse—having a considerable number of zero or nearly zero entries (Fig. 1c)—then an accurate reconstruction of the signal is still possible via a sparse recovery scheme [13].

The canonical form of a sparse recovery problem can be stated as follows: given an underdetermined linear model $\mathbf{y} = \mathbf{Ax}$, where $\mathbf{A} \in \mathbb{R}^{n \times p}$ is the so-called dictionary matrix ($n < p$), and $\mathbf{y} \in \mathbb{R}^{n \times 1}$ is the vector consisting of n scalar observations, find the unknown sparse signal $\mathbf{x} \in \mathbb{R}^{p \times 1}$ such that

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \qquad \text{s.t.} \quad \mathbf{y} = \mathbf{Ax}, \tag{2}$$

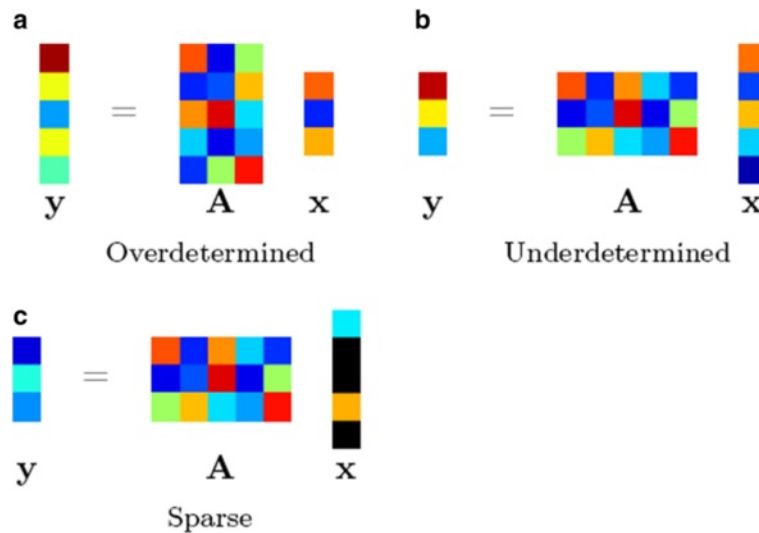**Fig. 1** Stylized visualization of three examples of linear equation systems. **A** and **y** represent the design matrix and observations, respectively; **x** is the unknown signal to be recovered. Positive and negative values are shown as *coloured blocks*, and zero entries are represented by *black blocks*. **a** Overdetermined system, where there are more observations (5) than unknowns (3). **b** Underdetermined system, where there are more unknowns (5) than observations (3). The signal **x** cannot be uniquely determined from such a system. **c** Underdetermined system with sparse signal. It is possible to recover **x** using sparse recovery methods as long as we know that **x** is sparse, even though the system is underdetermined and the exact positions of non-zero entries are not known *a priori*

where $\|\cdot\|_0$ represents ℓ_0 pseudo-norm, the number of non-zero entries.

Equation 2 is generally a non-deterministic polynomial-time (NP)-hard combinatorial problem [14]. In practice, it is more feasible to solve a relaxed form. We will briefly discuss various alternative formulations and corresponding solvers in Section 2.2.

1.2 Photometric stereo and sparse recovery

PST is often formulated as an overdetermined regression problem. The classical PST adopts three lights (hence three observations of luminance at each pixel location) [15] to solve for the 3D normal vectors. Later methods use more lights ranging from four to hundreds [6, 10, 16, 17]. Recently, a few attempts have been made to represent PST as an underdetermined system; firstly, in the case of calibrated lighting directions [1, 2], as addressed here, as well as for the alternative case of unknown lighting conditions [18, 19], not studied in this report. Reconfiguring PST as an underdetermined system means explicitly modelling the non-Lambertian error for each observation as additional unknowns. Suppose there are n lights (hence n equations for each pixel): the per-pixel number of unknowns would be $n + 3$ (three normal vector components and n non-Lambertian errors). As was already pointed out, such a system cannot be unambiguously solved through ordinary LS. Fortunately, if we make an assumption that the majority of luminance observations are approximately Lambertian, then the error vector is essentially a sparse vector with a large number of zero or approximately zero entries. Now that we have a sparse representation of the PST problem, we can solve it using a sparse recovery algorithm.

It has been shown by Wu et al. [1] and Ikehata et al. [2] that sparse PST behaves significantly more robustly than the classical PST method. However, the accuracy is contingent on the solver. At present, the most accurate solver for the sparse formulation is sparse Bayesian learning (SBL) as tested by Ikehata et al. in [2]. In the current study, we employ a modified form of the sparse representation given in [2], but solve it via a different approach—greedy sparse recovery algorithms.

1.3 Novel contributions

The main contributions of the current study are threefold:

1. We propose an alternative sparse formulation for PST (Eqs. 10 and 11). In previous PST studies [1, 2], the surface normal vector and the error vector are treated as two entities and are solved independently. In this study, we convert their formulation into a new canonical form of the sparse recovery problem by combining the two vectors into one large vector. Although such a "stacked" formulation is not novel

(e.g. [20]), it is used in the context of surface normal estimation for the first time. The advantage of this formulation is that it allows for a large repertoire of existing sparse recovery algorithms to be more straightforwardly applied to the PST problem.

2. We apply a greedy algorithm called orthogonal matching pursuit (OMP) [21–23], from information theory, to solve the PST problem. It has been previously demonstrated in [1, 2] that PST can be solved by several sparse recovery algorithms that fall into different categories, including augmented Lagrangian rank-minimization [1], ℓ_1 optimization approaches and probability-based methods [2]. However, the possibility of applying greedy solvers, an important category of sparse recovery algorithms, to the PST problem has never been explored. To the best of our knowledge, this study is the first of such attempt at employing greedy approaches to estimate surface normals within the framework of PST.

3. We numerically compare the performance of several normal vector recovery methods. Most notably, it is the first time that the normal estimation accuracy of our previously proposed method—LMS—has been tested and quantitatively demonstrated on complex models.

1.4 Overview

This paper is organized as follows: Section 2 provides a short survey on recent robust PST and sparse recovery methods. In Section 3, we provide a detailed description of our sparse formulation and the OMP algorithm. Experimental results and discussions are presented in Section 4, followed by several possible future research directions discussed in Section 5.

2 Related work

2.1 Robust photometric stereo

This section presents a brief overview of current PST methods. Since the original non-robust Lambertian-based PST [15], many methods have been proposed in an attempt to address non-Lambertian effects such as specularities and shadows. These approaches usually adopt a robust statistical method and/or an improved non-Lambertian reflectance model.

2.1.1 Statistics-based methods

In statistics-based methods, a robust statistical algorithm is employed to detect the non-Lambertian observations as outliers and exclude them from the estimation process in order to minimize their influence on the final result. Early examples include a four-light PST approach in which the values yielding significantly differing albedos are excluded [16, 24, 25]. In a similar five-light PST method [17], the highest and the lowest values, presumably corresponding

to highlights and shadows, are simply discarded. Another four-light method [26] explicitly included ambient illumination and surface integrability and adopted an iterative strategy, using current surface estimates to accept or reject each additional light based on a threshold indicating a shadowed value. The problem with these methods is that they rely on throwing away a small number of outlier observation values, whereas our robust sparse methods in the current study reaches the solution based on all observations, by correcting the non-Lambertian error of the outlier observations.

Willems et al. [27] used an iterative method to estimate normals. Initially, the pixel values within a certain range (10–240 out of 255) were used to estimate an initial normal map. In each of the following iterations, error residuals of normals for all lighting directions are computed and the normals are updated based only on those directions with small residuals. Sun et al. [28] showed that at least six light sources are needed to guarantee that every location on the surface is illuminated by at least three lights. They proposed a decision algorithm to discard only doubtful pixels, rather than throwing away all pixel values that lie outside a certain range. However, the validity of their method is based on the assumption that out of the six values for each pixel, there is at most one highlight pixels and two shadowed pixels. Mallick et al. [29] introduced a method based on colour space transformation to separate specular and diffuse components. Holroyd et al. [30] exploited the symmetries in the 2D slices of bidirectional reflectance distribution function (BRDF) obtained at each pixel to recover surface normal and tangent vectors. Both [29] and [30] can be applied to a great variety of surface reflectance, but they do not provide enough focus on the robustness against shadow pixels. Julià et al. [31] utilized a factorization technique to decompose the luminance matrix into surface and light source matrices. They consider the shadow and highlight pixels as missing data, with the objective of reducing the influence of these pixels on the final result.

Some recent studies utilize probability models as a mechanism to incorporate the handling of shadows and highlights into the PST formulation. Tang et al. [32] model normal orientations and discontinuities with two coupled Markov random fields (MRF). They proposed a tensorial belief propagation method to solve the *maximum a posteriori* (MAP) problem in the Markov network. Chandraker et al. [33] formulate PST as a shadow labelling problem where the labels of each pixel's neighbours are taken into consideration, enforcing the smoothness of the shadowed region, and approximate the solution via a fast iterative graph-cut method. Another study [8] employs a maximum likelihood (ML) imaging model for PST. In this method, an inlier map modelled via MRF is included in the ML model. However, the initial values of the inlier

map would directly influence the final result, whereas our sparse method does not depend on the choice of any prior.

A few other studies employ random-sampling-based methods. Using three-light datasets, Mukaigawa et al. [34] adopt a random sample consensus (RANSAC)-based approach to iteratively select random groups of pixels from different regions of the image, and the sampled group whose pixels are all taken from diffuse regions are used to calculate the coefficients in the linear equation. RANSAC is also used in a multiview context [9] as a robust fitting approach to select the points on a certain 3D curve. Drew et al. [10, 12] and Zhang and Drew [11] employ a LMS method. Instead of taking samples from different regions on the image, they use a denser image set (50 lights) and sample only from the observations at each pixel location. Non-Lambertian observations are rejected as outliers and excluded from the following LS step. Based on [33], Miyazaki et al. [35] used a median filtering approach similar to LMS but also considering neighbouring pixels. Instead of taking random samples, they simply compare all the three combinations of observations, which is feasible for the small number of lights used in their study. Although guaranteeing a high statistical robustness, these methods are computationally heavy since they usually rely on a large number of samples to take effect.

2.1.2 Non-Lambertian reflectance modelling

Instead of statistically rejecting non-Lambertian effects as outliers, another way to minimize their negative influence on surface normal recovery is to incorporate a more sophisticated reflectance model to directly account for the non-Lambertian components.

Tagare and de Figueiredo [36] constructed an m-lobed reflectance map model to approximate diffuse non-Lambertian surface-light interactions. In [25], a Torrance-Sparrow model is employed to estimate the roughness of the surface that is divided into different areas. Similarly, Nayar et al. [37] adopt a Torrance-Sparrow and Beckmann-Spizzichino hybrid reflectance model. Georghiades [38] applied Torrance-Sparrow model to handle the uncalibrated photometric stereo problem. Other mathematical models to encode surface reflectance include polynomial texture mapping (PTM) [39] and spherical harmonics (SH) [40]. Drew et al. [10] proposed a radial basis function (RBF) interpolation to handle the rendering of specularities and shadows.

Other studies use reference objects to facilitate the estimation of surface properties. In [41], an object with simple, known 3D geometry and approximately Lambertian reflectance (for instance, a white matte sphere) is present in the captured images. A look-up table is established that relates luminance observations at each pixel location and the surface orientation. Then, the surface properties

of other objects with similar reflectance as the reference object can simply be inferred from the look-up table. This method, however, only applies to isotropic materials. Hertzmann and Seitz [5] later revisited the idea of including reference material. By adopting an orientation-consistency cue assumption that two points on the surface with the same orientation have the same observed light intensity, they effectively cast PST as a stereoptic correspondence problem. This approach is capable of handling a wider range of anisotropic materials with a small number of reference objects, usually one or two. Similar to [5], an appearance-clustering method proposed by Koppal and Narasimhan [42], also adopting the orientation consistency cue, focuses on finding iso-normals across frames in a captured image sequence, and a classical PST approach may be applied later to obtain the accurate value of the surface normals. Although their method does not rely on the presence of a reference object, it does require the image sequence to be densely captured on a continuous path.

Recent studies attempt to solve a more complicated problem where neither shape nor material information of the object surface is available. Goldman et al. [43] employed an objective function that contains terms for both shape and material and proposed an iterative approach where the reflectance and shape are alternately optimized. The estimation of the material is an inseparable part of the reconstruction process so an explicit reference object is no longer needed. Alldrin et al. [6] also adopt a similar iterative approach that updates shapes and materials alternately. Their formulation is non-parametric and data-driven, and as such is capable of capturing an even wider range of reflectance materials. Ackermann et al. [7] proposed an example-based multi-view PST method which uses the captured object's own geometry as reference.

Yang et al. [44] include a dichromatic reflection model into PST for both estimating surface normals as well as separating the diffuse and specular components, based on a surface chromaticity invariant. Their method is able to reduce the specular effect even when the specular-free observability assumption (that is, each pixel is diffuse in at least one input image) is violated. However, this method does not address shadows and fails on surfaces that mix their own colours into the reflected highlights, such as metallic materials. Moreover, their method also requires knowledge of the lighting chromaticity—they suggest a simple white-patch estimator—whereas in our method, we have no such requirement. Kherada et al. [45] proposed a component-based mapping (CBM) method. They decompose the captured images into direct components (single bounce of light from a surface) and global components (illumination onto a point that is interreflected from all other points in the scene). They then model matte, shadow and specularity separately within each component. This method depends on a training phase, requires accurate disambiguation of direct and global contributions and has a high computational load. Shi et al. [46] introduced a bi-polynomial representation to model the low-frequency component of reflectance and used only the low-frequency information to recover shape and estimate reflectance.

The problem with these methods is that they usually do not work well against non-Lambertian effects that are not accounted for by the surface reflectance alone, such as cast shadows. In our current sparse method, we make no assumption of the surface reflectance property and treat all non-Lambertian effects (specularity and shadow) equally.

2.1.3 Sparse formulation

Recently, a few studies began to adopt sparse representation into PST. Wu et al. [1] model the matrix of all luminance observations as a linear combination of Lambertian and non-Lambertian components and represent the non-Lambertian error as an additive sparse noise vector. Under the assumption that most pixel observations approximately follow the Lambertian reflectance model, they obtain the solution by finding a sparse vector such that the rank of the Lambertian component matrix is minimized. The formulation is known as robust principal component analysis (R-PCA) in the field of sparse recovery. Specifically, they adopted a fast and scalable algorithm suitable for handling a large amount of data points, i.e. the augmented Lagrange multiplier method [47]. However, this method requires a shadow mask to be specified explicitly. Later, Ikehata et al. [2] reconsider PST as a pure sparse regression problem and aim to minimize the number of entries (i.e. the ℓ_0 pseudo norm) in the error matrix. They also add an ℓ_2 relaxation term to account for cases when the sparse assumption is violated. In order to avoid the difficult combinatorial problem involved in the minimization of ℓ_0 norm, they introduced two possible algorithms. One is to relax the ℓ_0 pseudo norm into ℓ_1 norm, as justified in [13, 48], and the solution is obtained via iteratively reweighted L1 minimization (IRL1) [49]. The other method is a hierarchical Bayesian approach called SBL [50]. It has been shown that SBL has an improved accuracy over IRL1 at the expense of lower efficiency, and both IRL1 and SBL perform better than R-PCA [2]. Independently of [1, 2], a similar, but quite complex, schema called alternating direction method of multipliers (ADMM) is developed in a recent paper by Adler et al. [51]. That paper in part applies sparse coding to PST and thus provides a motivating source for the present paper; however, our paper makes use of a radically different formulation as well as ties the PST problem into well-studied sparse solver algorithms. Moreover, that paper is aimed at producing

a matte image with specularity and shadows attenuated, whereas the present work is aimed at accurate normal-vector recovery. Whereas our use of OMP uses a greedy algorithm where each component is picked one at a time, in contrast, the ADMM approach adjusts all the components in each iteration. The present study is the first to use greedy approaches to surface normal estimation within a PST framework.

Sparse methods have also found their use in uncalibrated PST, where the lighting directions are not known (but note that in this study we do assume known lighting directions so that these works are somewhat peripheral). Favaro et al. [18] incorporate the rank-minimization algorithm proposed in [1] into the uncalibrated PST problem as a pre-processing step to remove shadow and specularity effects. Argyriou et al. [19] recently also adopt a sparse representation framework to decide the weights for finding the best illuminants to use, again with the lighting directions unknown.

2.2 Sparse recovery methods

As was pointed out in Section 1.1, the canonical form of the sparse recovery problem (Eq. 2) is NP-hard [14] and cannot be solved efficiently as-is. In this section, we summarize alternative formulations to Eq. 2 and several types of solvers.

The first type of approach is convex ℓ_1 relaxation. It has been shown that for a dictionary matrix $\mathbf{A}$ that satisfies a certain restriction, Eq. 2 is likely to be equivalent to an ℓ_1 minimization problem [13, 48]:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \qquad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{3}$$

which can be solved via convex optimization techniques such as interior-point (IP) methods [52], gradient projection [53], IRL1 [49] and so forth.

Alternatively, sparse recovery can be achieved via greedy algorithms. The basic idea of such an algorithm is employing an iterative method to find the collection of non-zero entries, or *support*, of the signal $\mathbf{x}$, and then recover $\mathbf{x}$ via LS using only the observations in the support.

One of the most notable greedy algorithms is OMP [21–23], an improvement over the simple matching pursuit (MP) algorithm [54]. In OMP, a column $\mathbf{a}_j$ in $\mathbf{A}$ is iteratively chosen such that $\mathbf{a}_j$ is most greatly correlated with the current residual $\mathbf{r}$. Then, $\mathbf{r}$ is updated by taking into consideration the contribution of $\mathbf{a}_j$. The algorithm is terminated as a fixed number of non-zero entries are recovered or other stopping criteria are met. Then, a simple LS is performed only on a submatrix of $\mathbf{A}$ consisting of the columns chosen by OMP, and the regressed result will be assigned only to the signal entries corresponding to the selected columns. The columns that are not selected by OMP, on the other hand, will not be used in the final LS

step, and their corresponding signal entries are simply set to zero.

In fact, OMP approximately solves the following k-sparse recovery problem:

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \qquad \text{s.t.} \quad \|\mathbf{x}\|_0 \le k. \tag{4}$$

Many state-of-the-art greedy algorithms nowadays are based on OMP. Examples include regularized OMP (ROMP) [55, 56], stagewise OMP (StOMP) [57], compressive sampling matching pursuit (CoSaMP) [58], probability OMP (PrOMP) [59], look ahead OMP [60], OMP with replacement (OMPR) [61], A* OMP [62] etc.

Another type of solvers employ a thresholding step to iteratively refine the recovered support, i.e. the selection/rejection of an entry at each step, is decided by whether the value of a certain function dependent on this entry falls below a given threshold. Algorithms in this category include iterative hard thresholding (IHT) [63], subspace pursuit (SP) [64], approximate message passing (AMP) [65], two-stage thresholding (TST) [66], algebraic pursuit (ALPS) [67] etc.

The fourth category is probability-based algorithms. These methods assume the signal to be recovered follows a specific probability distribution and solve the sparse recovery problem with statistical methods such as ML or MAP estimation. SBL [50] is one of the major algorithms in this category and has already been applied in the context of PST [2].

3 Sparse regression
3.1 Sparse formulation for photometric stereo
In this section, we explore the possibility of formulating and solving PST as a sparse regression problem. Since only the normal recovery is studied in this paper, we omit the albedo α from all equations in this and the following sections for simplicity and always use $\mathbf{n}$ to represent the unnormalised surface normal vector unless otherwise specified.

Here, we assume a Lambertian reflectance model with an additional term $e \in \mathbb{R}$ to account for the non-Lambertian error. Hence, the observed luminance y can be expressed as:

$$y = \boldsymbol{l} \cdot \mathbf{n} + e, \tag{5}$$

where $\boldsymbol{l} \in \mathbb{R}^3$ and $\mathbf{n} \in \mathbb{R}^3$ represent the lighting direction and surface normal, respectively. For each pixel, we have n observations $\mathbf{y} = (y_1, y_2, \ldots y_n)^T \in \mathbb{R}^n$. Now, let us write Eq. 5 in vector form

$$\mathbf{y} = \mathbf{L}\mathbf{n} + \mathbf{e}, \tag{6}$$

where $\mathbf{L} = (\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots \boldsymbol{l}_n)^T \in \mathbb{R}^{n \times 3}$ and $\mathbf{e} = (e_1, e_2, \ldots e_n)^T \in \mathbb{R}^n$.

Equation 6, containing n linear equations but $n + 3$ unknowns (n components in $\mathbf{e}$ and three components in

n), is effectively an underdetermined problem and as such cannot be solved unambiguously. However, if the error **e** is a sparse matrix, i.e. most or at least a great percentage of its elements are zero, then it is still possible to recover **e** exactly or almost exactly by solving the following sparse regression problem:

$$\min_{\mathbf{n,e}} \|\mathbf{e}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{Ln} + \mathbf{e}. \tag{7}$$

In Eq. 7, $\|\cdot\|_0$ represents the ℓ_0 pseudo-norm or the number of non-zero elements in **e**. This formulation, however, has two major issues: (1) it is an NP-hard combinatorial problem and (2) real-world scenes may contain a large variety of materials that are only poorly approximated by the Lambertian reflectance model. For those materials, it is very likely that **e** is not strictly sparse. Thus, the equality constraint is very hard to be satisfied. Instead, it is more realistic to use an inequality constraint with a user-defined error tolerance ϵ

$$\min_{\mathbf{n,e}} \|\mathbf{e}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{Ln} - \mathbf{e}\|_2 \leq \epsilon. \tag{8}$$

Alternatively, if we care more about how much the reconstructed luminance approximates real observation rather than the sparsity of **e**, then it would be more natural to reformulate Eq. 8 as

$$\min_{\mathbf{n,e}} \|\mathbf{y} - \mathbf{Ln} - \mathbf{e}\|_2 \quad \text{s.t.} \quad \|\mathbf{e}\|_0 \leq s, \tag{9}$$

where the scalar s is the sparsity of vector **e**. To further simplify Eq. 9, we propose merging **n** and **e** into one large vector and treating them as one entity, i.e.

$$\begin{aligned}\mathbf{y} &= \mathbf{Ln} + \mathbf{e} \\ &= \mathbf{Ln} + \mathbf{Ie} \\ &= (\mathbf{L, I}) \begin{pmatrix} \mathbf{n} \\ \mathbf{e} \end{pmatrix} \\ &= \mathbf{Ax}, \end{aligned} \tag{10}$$

where $\mathbf{I} \in \mathbb{R}^{n \times n}$ is an $n \times n$ identity matrix, $\mathbf{A} = (\mathbf{L, I}) \in \mathbb{R}^{n \times (n+3)}$ is a new merged dictionary matrix and $\mathbf{x} = (\mathbf{n^T, e^T})^T \in \mathbb{R}^{(n+3) \times 1}$ is the combined vector of all the unknown variables. Hence, Eq. 9 can be rewritten as

$$\min_{\mathbf{n,e}} \|\mathbf{y} - \mathbf{Ax}\|_2 \quad \text{s.t.} \quad \|\mathbf{x}\|_0 \leq s. \tag{11}$$

The stacked formulation was inspired by the work of Wright et al. [20, Eq. 20]. However, in [20], both the signal and the noise are assumed sparse, whereas in our case, the signal (normal vector) has only three components and is not at all sparse.

By formulating our problem in the form of Eq. 11, we can now take advantage of existing algorithms to efficiently achieve an accurate solution. One such solver is a greedy algorithm known as OMP [21–23], which is known for its high accuracy and low time-complexity. We will describe this algorithm in Section 3.2 in detail.
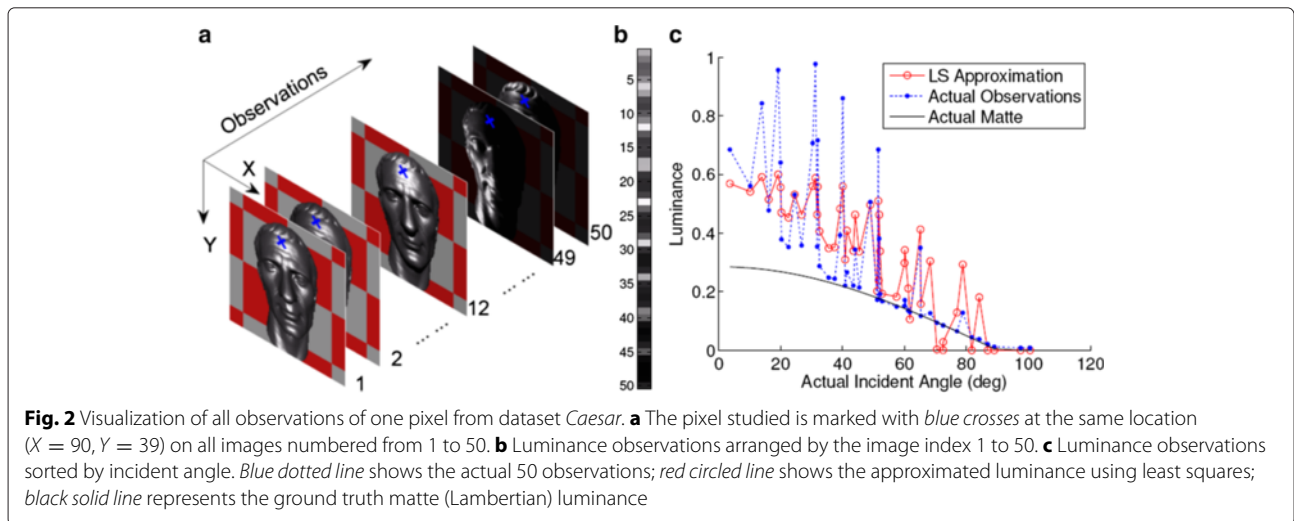
Previously, Ikehata et al. [2] proposed a different formulation to Eq. 11. They expressed the PST problem in a so-called Lagrangian form, i.e.

$$\min_{\mathbf{n,e}} \|\mathbf{y} - \mathbf{Ln} - \mathbf{e}\|_2^2 + \lambda \|\mathbf{e}\|_1 \tag{12}$$

and applied two solving algorithms: IRL1 minimization and SBL. They showed that SBL provides a more accurate estimation but is more computationally expensive. Later, in Section 4, we will show that our OMP solver produces a more accurate result than SBL with comparable efficiency to IRL1.

3.2 Orthogonal matching pursuit

Sparse recovery problems like Eq. 11 can be solved via many different methods (see Section 2.2 for a brief overview). Here, we choose to apply the classical greedy OMP to our surface normal recovery problem. Given the
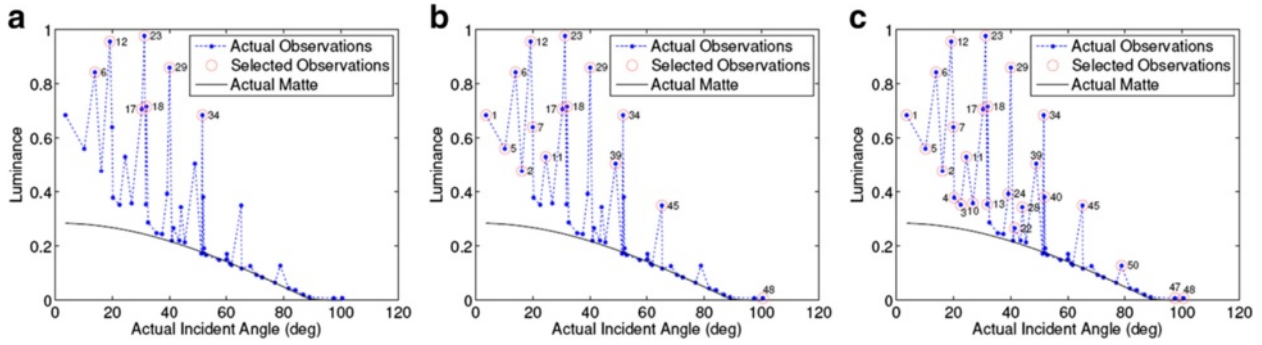

Fig. 2 Visualization of all observations of one pixel from dataset *Caesar*. **a** The pixel studied is marked with *blue crosses* at the same location ($X = 90$, $Y = 39$) on all images numbered from 1 to 50. **b** Luminance observations arranged by the image index 1 to 50. **c** Luminance observations sorted by incident angle. *Blue dotted line* shows the actual 50 observations; *red circled line* shows the approximated luminance using least squares; *black solid line* represents the ground truth matte (Lambertian) luminance

Fig. 3 Outliers identified by orthogonal matching pursuit. **a** Outliers with great non-Lambertian error (*red circles*) detected in iterations 4–10. **b** Outliers with medium to great error (*red circles*) detected in iterations 4–18. **c** All outliers (*red circles*) detected as of iteration 28. *Blue dotted lines* show actual luminance observations in all three plots

linear model in Eq. 10, the basic idea of OMP is to iteratively select columns of the dictionary matrix $\mathbf{A}$ that are most closely correlated with the current residuals, then project the observation $\mathbf{y}$ to the linear subspace spanned by the columns selected until the current iteration. We denote each column of $\mathbf{A}$ as $\mathbf{A}_j$.

Let i be the current number of iterations and $\mathbf{r}_i$ and c_i the residuals and the subset of selected columns in $\mathbf{A}$ at the ith iteration, respectively. Let $\mathbf{A}(c_i)$ and $\mathbf{x}(c_i)$ represent the columns indexed by c_i in $\mathbf{A}$ and the entries indexed by c_i in the signal $\mathbf{x}$ to be recovered, respectively. The OMP algorithm [23], as we here apply to our PST problem, can be summarized as follows:

Algorithm Orthogonal matching pursuit

1. Normalise each column of the dictionary matrix $\mathbf{A}$ and denote the resulting matrix as $\mathbf{A}'$, i.e. $\|\mathbf{A}'_j\|_2 = 1$ for $j = 1, 2, \ldots, p$. Initialize the iteration counter $i = 1$, residual $\mathbf{r}_0 = \mathbf{y}$ and $c_0 = \varnothing$.
2. Find a column $\mathbf{A}'_t$ ($t \in \{1, 2, \ldots, p\} - c_{i-1}$) that is most closely correlated with the current residual. Equivalently, solve the following maximization problem:

$$t = \underset{j}{\operatorname{argmax}} \|\mathbf{A}'^T_j \mathbf{r}_{i-1}\| \quad (13)$$

3. Add t to the selected set of columns, i.e. update $c_i = c_{i-1} \cup t$, and use $\mathbf{A}'(c_i)$ as the current selected subset of $\mathbf{A}'$.
4. Project the observation $\mathbf{y}$ onto the linear space spanned by $\mathbf{A}'(c_i)$. The projection matrix is calculated as follows:

$$\mathbf{P} = \mathbf{A}'(c_i)(\mathbf{A}'(c_i)^T \mathbf{A}'(c_i))^{(-1)} \mathbf{A}'(c_i)^T. \quad (14)$$

5. Update the residuals with respect to the new projected observation

$$\mathbf{r}_i = \mathbf{y} - \mathbf{P}\mathbf{y}. \quad (15)$$

6. Increment i by 1. If $i > n/2 + 3$ (= 28 for our typical datasets of 50 images), then proceed to step 7, otherwise go back to step 2.
7. Solve only for the entries indexed by c_i in signal $\mathbf{x}$ using the original, **unnormalised** design matrix $\mathbf{A}$, and simply set the rest of the entries to 0, i.e.

$$\mathbf{x}(c_i) = \mathbf{A}(c_i)^{\dagger} \mathbf{y} \quad (16)$$

and

$$x_j = 0 \qquad \text{for each} \quad j \notin c_i. \quad (17)$$

8. Take the first three entries in $\mathbf{x}$ as the solutions for the x, y and z component of the normal vector, respectively,
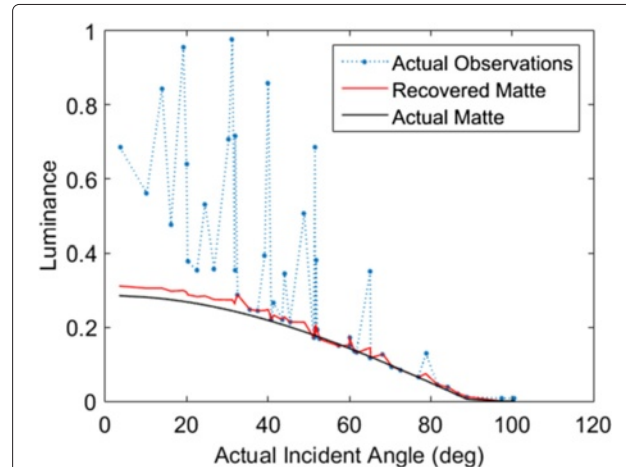
$$\mathbf{n} = (x_1, x_2, x_3). \quad (18)$$

Fig. 4 Matte component recovered by orthogonal matching pursuit. *Blue dotted line* represents actual luminance observations; *black solid line* shows the actual matte model; *red solid line* shows the matte recovered by OMP

In our formulation Eq. 10, we merge the normal and the errors into a large vector, so the components of the two vectors are treated equally by OMP. In each iteration, which column in the dictionary matrix is to be chosen purely depends on its correlation with the current residuals. Thus, there is no strict mathematical guarantee that the normal vector components will be selected in the first s iterations. Indeed, this failure could happen if the non-Lambertian error vector accounts for most of the observations. However, since the observed luminance is usually a function of the surface normal, it is expected that the normal vector components are more closely correlated to the observations than the sporadic non-Lambertian errors. In our experiments, all normal vector components are usually selected within the first few iterations (<10). On the other hand, if one or two components of the surface normal are rather small, then they might not be selected by our algorithm. However, since they are very close to zero anyway, simply treating them as zero would not negatively impact the accuracy of our estimation.

One of the biggest advantages of OMP is its low computational cost and straightforward implementation. We have found that it is significantly faster than LMS as well as other state-of-the-art robust regression methods that
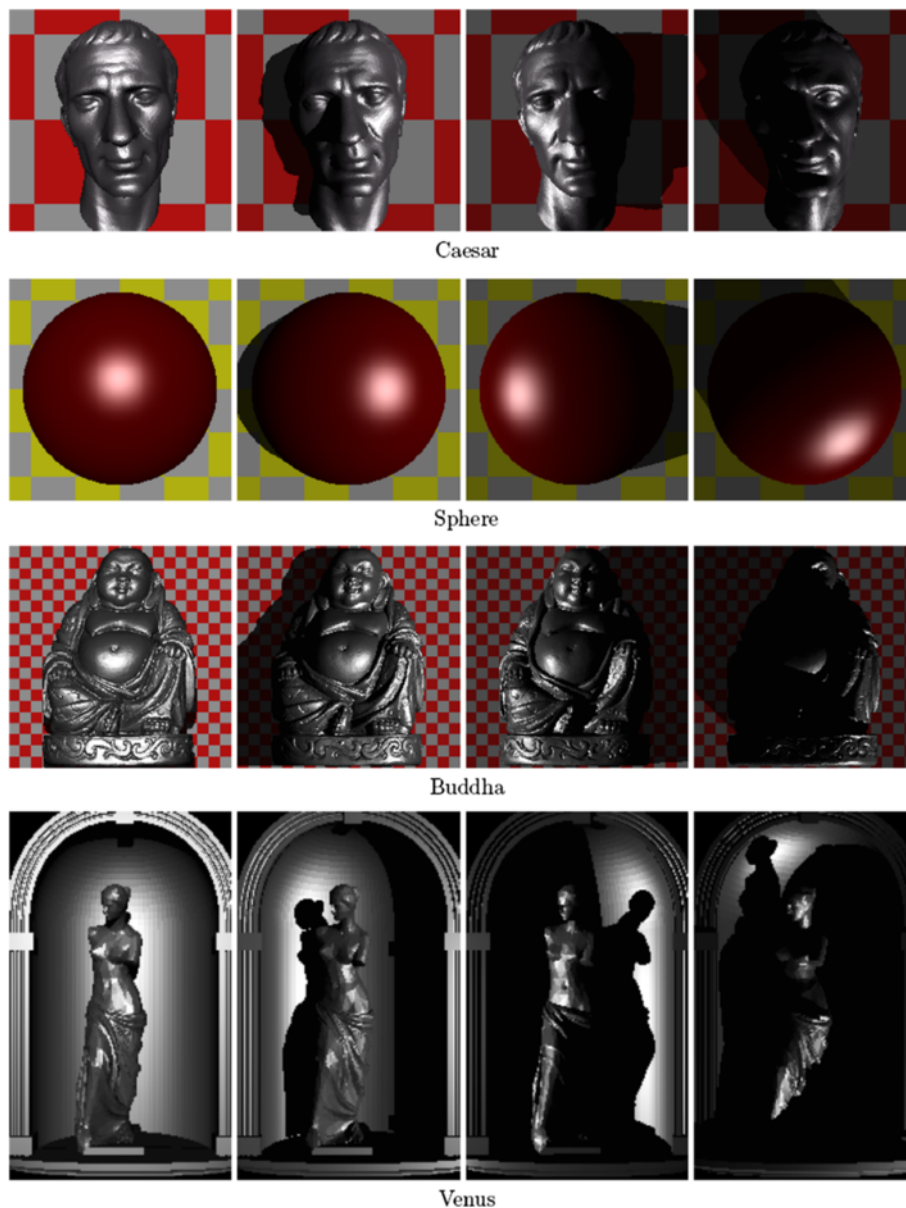

Fig. 5 Sample images from four synthesized datasets rendered with POV-Ray. From top row to bottom row: *Caesar, Sphere, Buddha* and *Venus*

have been applied in the context of PST (see Section 4.3 and Fig. 21 for more details). Note that for our particular choice of the design matrix **A**, the correlation between any column in the identity matrix and the residual **r** can be simply represented by one element in **r**. Therefore, the inner product in Eq. 13 may be reduced to finding the maximum entry in **r**. This observation allows for an even more efficient implementation. In this work, however, we still implement OMP according to Eq. 13 for generality.

3.2.1 Normalization and orthogonality

As a requirement of the standard OMP algorithm, we used the column normalised version of the design matrix **A**′ in the column selecting process. After normalization, the first three columns in **A**′ no longer hold the correct value of lighting vector components. In other words, it appears that the lighting directions are modified by normalization. However, this observation does not negatively affect our results. In step 2 of OMP, the column most correlated with the current residual vector is selected. Normalization only makes sure one column does not have a numerical advantage over another simply because it has a greater L2 norm. Therefore, normalization does not interferes with the selection of the outliers. On the contrary, it enforces the correctness of selection. It is also important to note that after the outliers are selected, we use the original unnormalised dictionary matrix **A**, instead of **A**′, to make sure that the normal vector are recovered on the actual lighting directions.
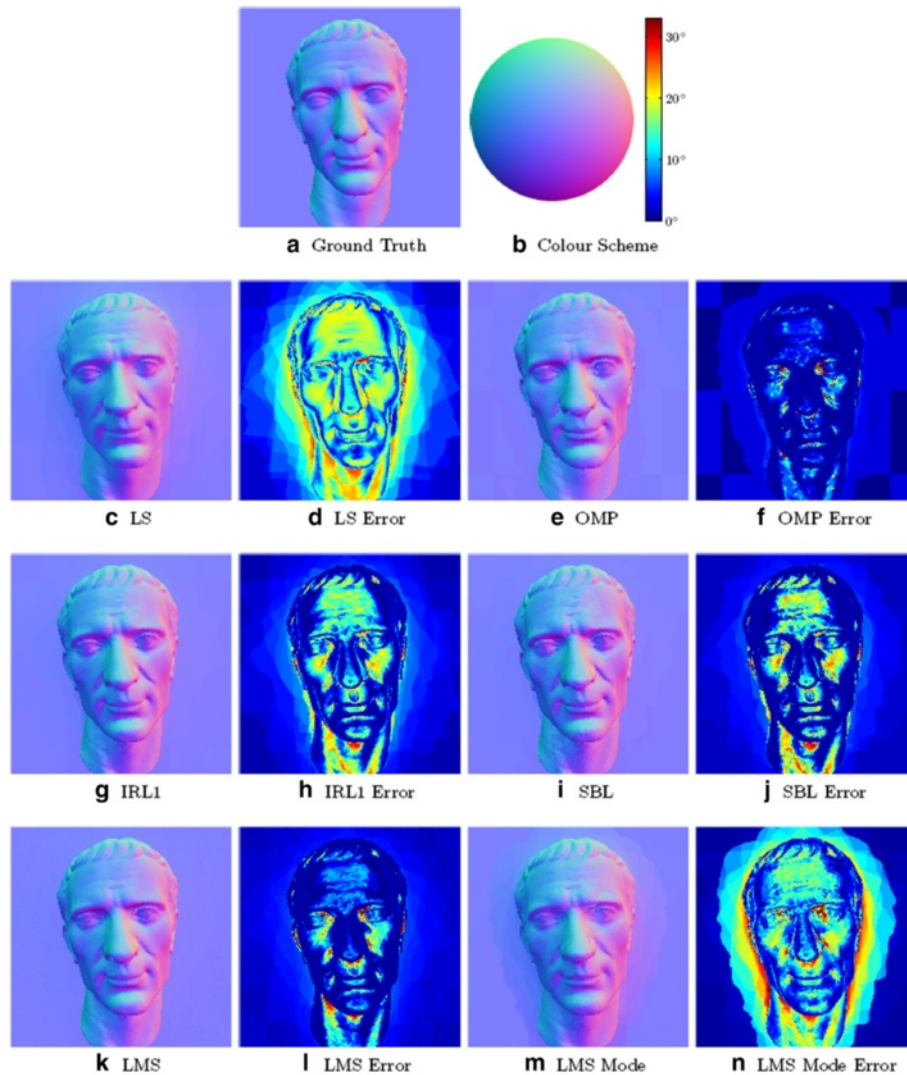

Fig. 6 Normal maps of a head statue of Caesar recovered using various methods. **a** Ground truth normal map; **b** colour wheel and colour bar used for normal and angular error visualization, respectively. Angular error is measured in degrees. **c**, **e**, **g**, **i**, **k**, **m** Normal map recovered using LS, OMP, IRL1, SBL, LMS and LMS mode, respectively. **d**, **f**, **h**, **j**, **l**, **n** Angular error of normal maps recovered using the aforementioned methods

Another issue worth noting is the orthogonality of the dictionary matrix **A**. It has been shown that if **A** satisfies a restricted isometry property (RIP), then the exact recovery of signal **x** may be possible [68, 69]. Essentially, RIP specifies a near-orthonormal condition for **A**. Although our dictionary matrix **A** unfortunately does not obey the RIP property in its general form, we still argue that this matrix is near-orthogonal: with our uniform light distribution, the first three columns are indeed near-orthogonal (the dot products between column 1–2, 1–3 and 2–3 are 8.85×10^{-8}, -1.04×0^{-6} and -5.17×10^{-7}, respectively). The rest of **A** is a large identity matrix **I**, which itself is orthonormal. Also, due to the large number of zeros in **I**, the dot product of any of the first three columns and any column in **I** is a rather small number ($10^{-3} - 10^{-2}$ scale on average). Thus, although it is yet to be strictly proven, we speculate that the dictionary matrix **A** in near-orthogonal enough for our purpose of recovering the three components of normal out of the 53-element signal. As our results show, OMP indeed achieves highly precise recovery of surface normal for most of the pixel locations (see Section 4). We also show that even with a very biased lighting distribution (such that the orthogonality of the first three column are greatly reduced), OMP still provides an accepted recovery with higher accuracy than other sparse methods (see Section 4.2.4).

3.2.2 Stopping criterion

For simplicity, here we set the stopping criterion as a fixed number (s) of iterations. We make a conservative
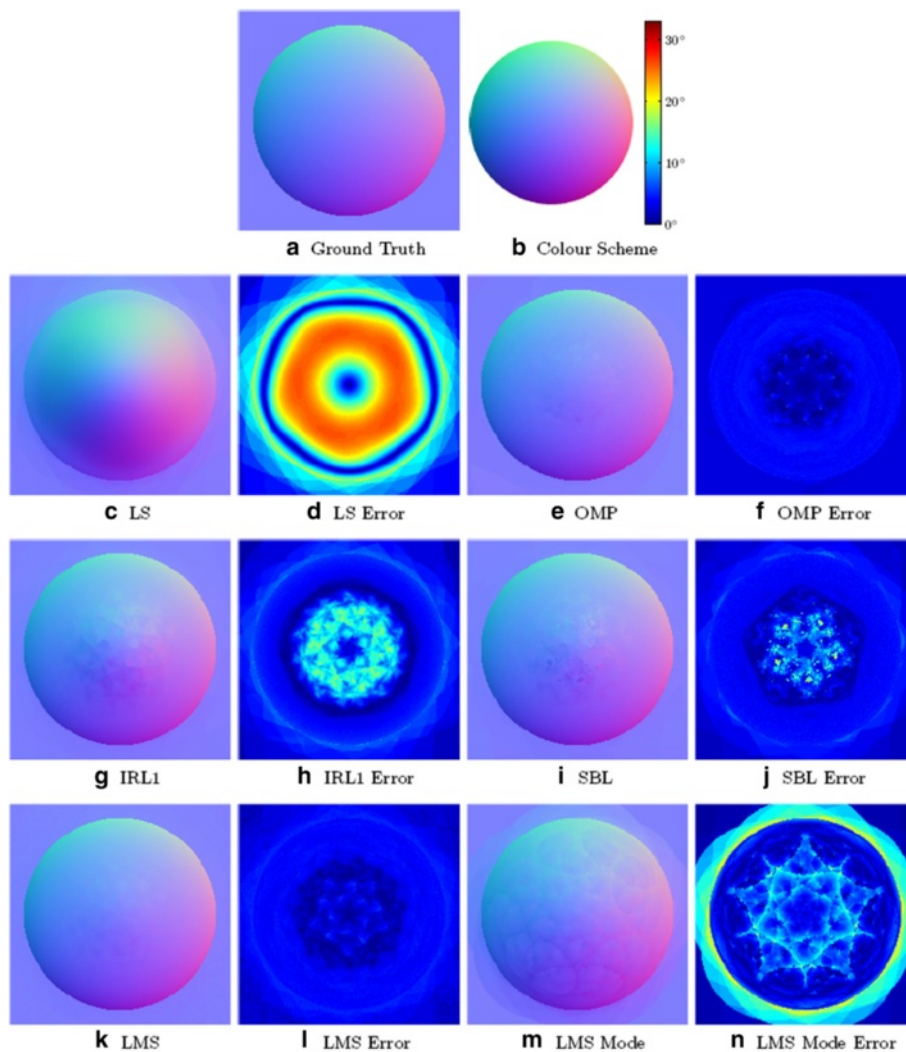

Fig. 7 Normal map of an ideal sphere recovered using various methods. **a** Ground truth normal map; **b** colour wheel and colour bar used for normal and angular error visualization, respectively. Angular error is measured in degrees. **c**,**e**, **g**, **i**, **k**, **m** Normal map recovered using LS, OMP, IRL1, SBL, LMS and LMS mode, respectively. **d**, **f**, **h**, **j**, **l**, **n** Angular error of normal maps recovered using the aforementioned methods

assumption that 50 % of the observed pixels are polluted by non-Lambertian noise. Thus, for our typical datasets of $n = 50$ images and normal vectors with three components, the stopping criterion is $i > s = n/2 + 3 = 28$. This criterion ensures that there is always a moderate number of observations (25) available for regression.

An alternative choice of stopping criterion is based on the residual, i.e. $|\mathbf{r}| <$ threshold. It has been proved that in the matching pursuit and its orthogonal version, OMP, are guaranteed to converge [21, 54]. Thus, such a stopping criterion is theoretically viable. Indeed, we have validated that in our OMP-based method, the residual converges on all pixels used in our datasets. However, we have also noticed in our tests that setting a hard threshold for all pixels results in a slightly decreased accuracy

than using the sparsity-based criterion (results not shown). Therefore, in the current study, we will continue using the sparsity-based criterion, i.e. $i >= 28$.

3.3 Visual demonstration

In this section, we demonstrate how OMP enforces robustness onto the normal recovery process by using a simple example. Particularly, we use a synthesized dataset *Caesar* (see Section 4.1.1 for more information) and study all the 50 observations of one pixel (marked by blue crosses in Fig. 2a) at location ($X = 90, Y = 39$) where the ground truth normal vector is $\mathbf{n}_{gt} = (-0.0780, 0.1828, 0.9801)$. The luminance profile of these observations, sorted by incident light angle, are shown in Fig. 2c (blue dotted line), along with the actual matte
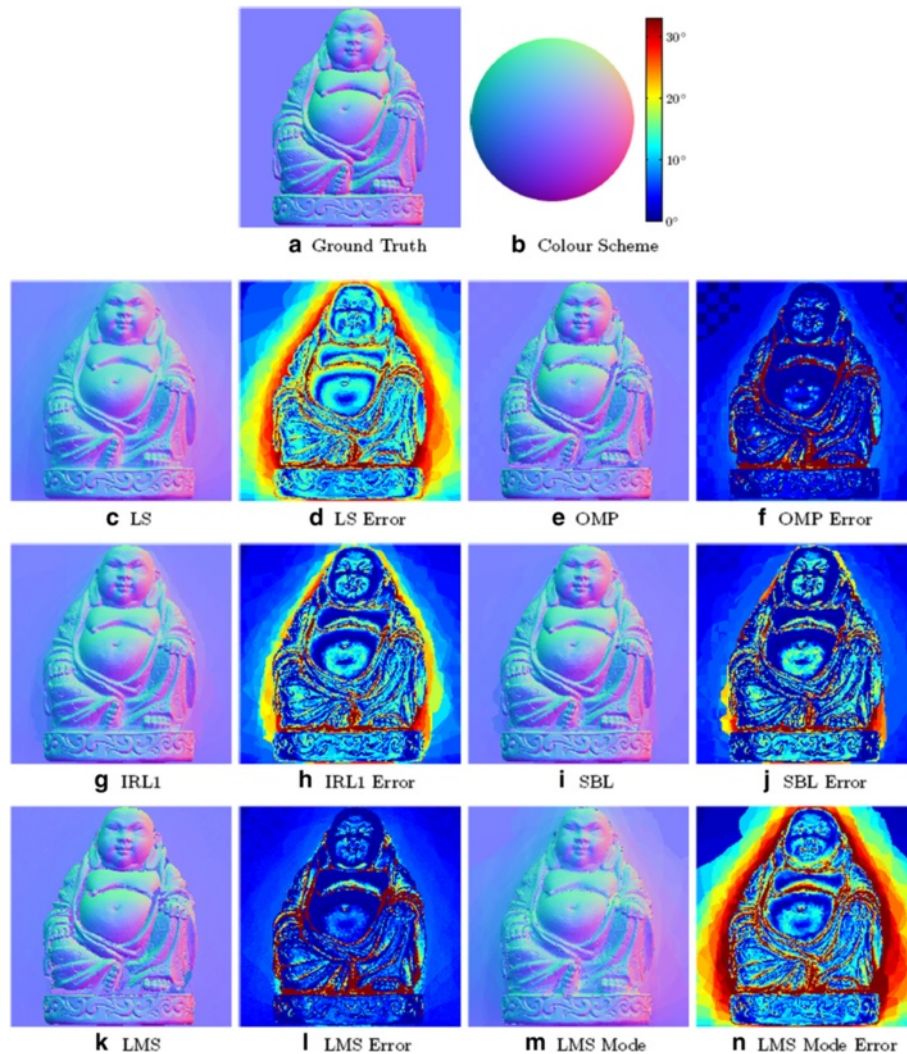

Fig. 8 Normal maps of a head statue of Buddha recovered using various methods. **a** Ground truth normal map; **b** colour wheel and colour bar used for normal and angular error visualization, respectively. Angular error is measured in degrees. **c**, **e**, **g**, **i**, **k**, **m** Normal map recovered using LS, OMP, IRL1, SBL, LMS and LMS mode, respectively. **d**, **f**, **h**, **j**, **l**, **n** Angular error of normal maps recovered using the aforementioned methods

model curve (black solid line), i.e. the theoretical values of luminance, if the surface is purely Lambertian. It is obvious from Fig. 2c that due to the existence of specular reflection, a good percentage of observations (especially when the incident angle is small) deviate from the values predicted by a matte model.

The naive LS regression, when applied to this pixel, attempts to approximate the values of all observations without taking the actual matte model into consideration (Fig. 2c, red line marked with circles). Naturally, the LS result $\mathbf{n}_{\mathrm{LS}} = (-0.1578, 0.4852, 0.8600)$ deviates greatly from the ground truth $(-0.0780, 0.1828, 0.9801)$.

On the other hand, OMP first attempts to identify s entries, one in each iteration, from the stacked signal $\mathbf{x} = (x_1, x_2, \ldots x_{n+3})^T \in \mathbb{R}^{n+3}$ (see Eq. 10). Usually, these s entries include three components for normal vectors (x_1, x_2, x_3) and ($s - 3$) components from the remaining n entries ($x_4, x_5, \ldots, x_{n+3}$) that correspond to error values. When the OMP algorithm as described in Section 3.2 is applied to this pixel, it behaves as follows:

Iterations 1–3: The entries that correspond to normal vectors, x_3, x_2 and x_1, are selected in the order listed. This is not a coincidence since the first three columns of the dictionary matrix $\mathbf{A}$ are overall more strongly correlated to the observations than any of the rest of the columns that correspond to noise values. Also, we noticed that these entries are in fact selected in order of the absolute value of their corresponding normal components. For instance, the third component of the ground truth normal $(-0.0780, 0.1828, 0.9801)$ is greater than the other two components. Therefore, x_3 gets selected in the first iteration.

Iterations 4–10: Entries x_{26}, x_{15}, x_{32}, x_9, x_{37}, x_{21} and x_{20} are selected sequentially. These entries correspond to non-Lambertian errors at observation #23, #12, …#17, respectively (marked with red circles in Fig. 3a). Note that the indices of observations mentioned here (23, 12, …) are equal to the entry indices found (26, 15, …) minus 3, since the first three elements in $\mathbf{x}$ do not represent errors. We notice that the corresponding observations of these selected entries all have very high error values. As in iterations 1–3, these error entries are also selected in order of their absolute value. For instance, observation #23 (incident angle $\approx 32°$) has the greatest non-Lambertian error; therefore, its corresponding error entry x_{26} is selected in iteration 4, before other entries.

Table 1 Statistics for the angular error between the normal maps recovered for various methods and the ground truth, for three synthesized datasets. All numbers are shown in degrees

	25 % Quantile	Median	Mean	75 % Quantile	STD
Dataset Caesar					
LS	4.669	8.586	9.672	14.32	5.875
LMS mode	1.313	6.769	8.148	13.46	7.620
LMS	1.335	2.692	3.700	4.305	4.242
IRL1	1.854	3.339	5.276	6.440	5.080
SBL	1.576	2.441	4.487	4.488	5.158
OMP	0.7571	2.624	2.424	2.690	3.332
Dataset Sphere					
LS	7.255	12.23	13.55	20.79	7.566
LMS mode	1.658	4.939	6.211	9.907	4.901
LMS	2.588	3.601	3.261	4.030	1.103
IRL1	3.045	4.253	4.960	5.662	3.119
SBL	2.762	3.620	3.671	4.239	1.822
OMP	2.617	3.051	3.179	3.991	0.9324
Dataset Buddha					
LS	9.039	14.03	15.29	20.19	8.25
LMS mode	7.304	14.70	17.42	25.03	13.79
LMS	1.985	4.736	8.454	6.391	15.35
IRL1	4.066	7.354	10.08	13.28	9.178
SBL	2.412	5.295	8.034	9.333	9.479
OMP	1.289	3.049	7.065	5.495	14.17

Iterations 11–18: Another eight entries x_4, x_{10}, x_{42}, x_8, x_{14}, x_{51}, x_{48} and x_5 are selected sequentially. Their corresponding observations have medium error values (Fig. 3b).

Iterations 19–28: Select the rest of the error entries x_{43}, x_{27}, x_{50}, x_{31}, x_{16}, x_7, x_{13}, x_6, x_{53} and x_{25}. The corresponding observations have small error values (Fig. 3c).

Through the 28 iterations above, we have obtained 28 indices; 3 of them correspond to the normal-vector components and the remaining 25 represent the observations that have significant non-Lambertian effect, i.e. non-zero values in signal $\mathbf{x}$ in the sparse regression problem $\mathbf{y} =$ $\mathbf{Ax}$ (Eq. 10). Suppose the indices of 25 selected non-Lambertian outliers are collectively represented as $c_{\text{out}} \subset \{1, 2, \ldots, 50\}$, we can obtain the normal vector $\mathbf{n}$ and an error vector $\mathbf{e}$ by solving the following equation (which is essentially the same as 16):

$$\mathbf{y} = (\mathbf{L}, \mathbf{I}(c_{\text{out}})) \begin{pmatrix} \mathbf{n} \\ \mathbf{e}(c_{\text{out}}) \end{pmatrix}. \tag{19}$$

For our sample pixel, the above equation gives $\mathbf{n}_{\text{OMP}} = (-0.0877, 0.2282, 0.9697)$, which well approximates the ground truth $\mathbf{n}_{\text{gt}} = (-0.078, 0.1828, 0.9801)$ compared to the naive LS result $\mathbf{n}_{\text{LS}} = (-0.1578, 0.4852, 0.8600)$.
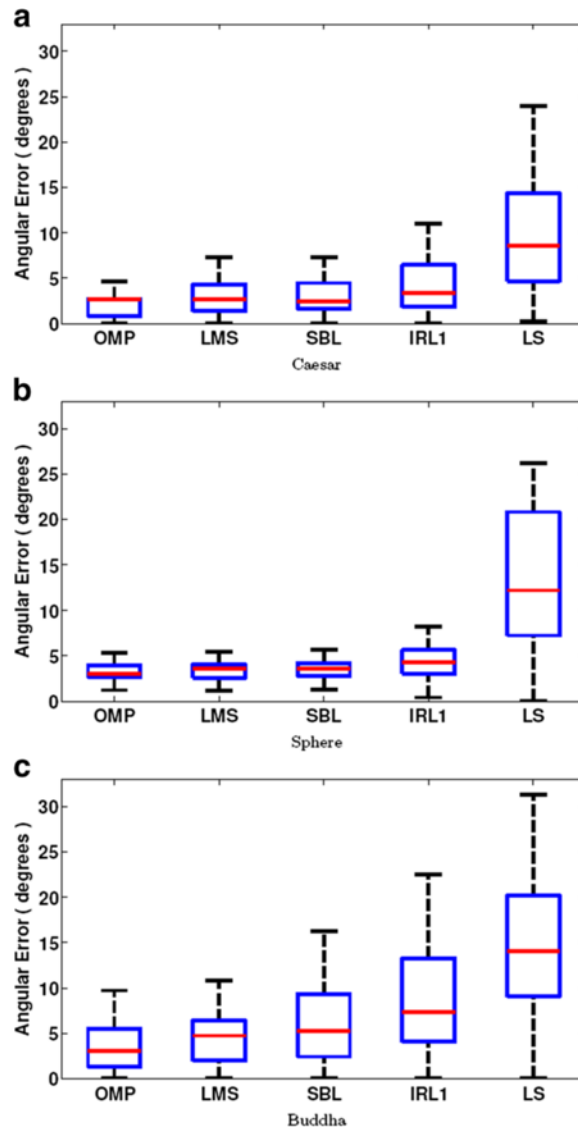

Fig. 9 Angular error of normal maps recovered using various methods (**a**–**c**). *Red horizontal lines* indicate medians. Upper and lower border of the *blue boxes* represent third (Q3) and first quartile (Q1), respectively. Upper and lower whiskers show 1.0 interquartile range (the difference between the upper and lower quartiles, the IQR) extended from Q3 and Q1, respectively. The whiskers show the quantile Q1 and Q3 extended by 1.0 IQR (as opposed to the typically shown 1.5 IQR)
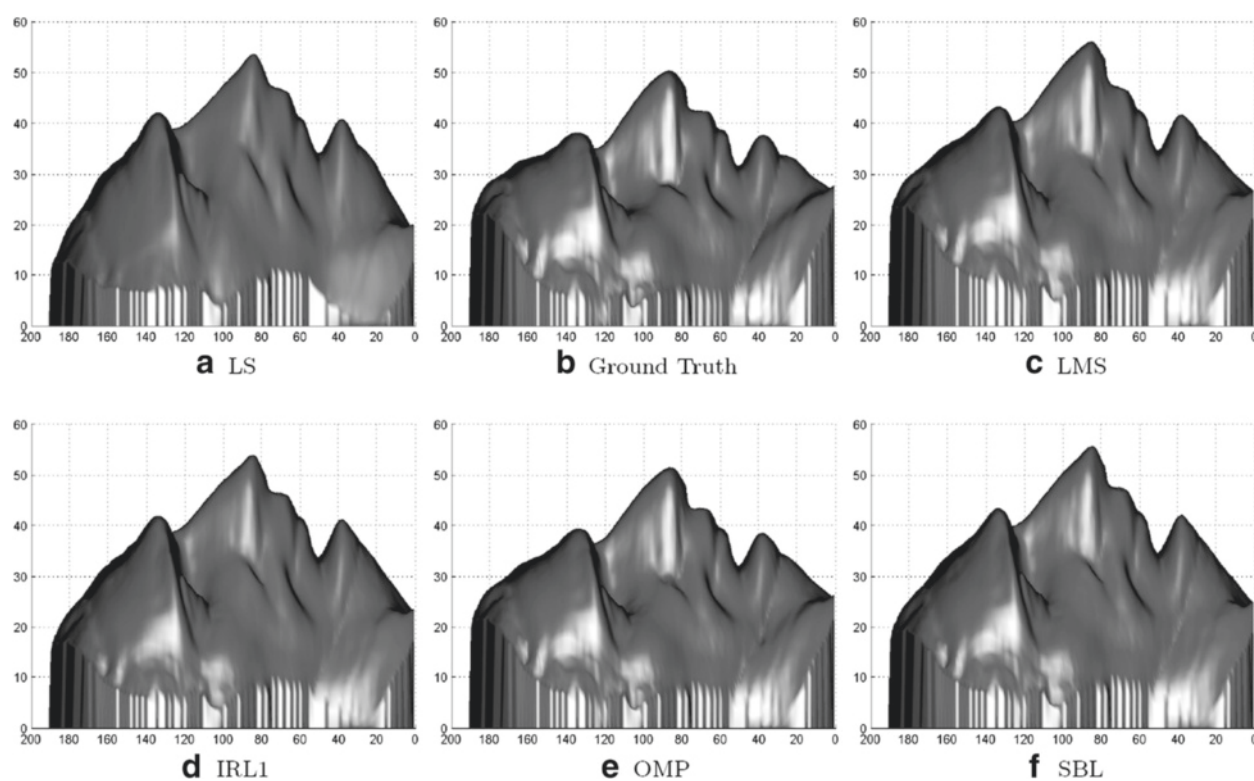
Fig. 10 Three-dimensional surfaces reconstructed from normal maps. **b** Depth map recovered with the ground truth normal map. **a**, **d**, **e**, **c**, **f** Depth maps recovered using LS, IRL1, OMP, LMS and SBL, respectively

Fig. 11 Sample images from three datasets of real-world scenes. From top row to bottom row: *Gold*, *Elba* and *Frag*

Additionally, we can directly recover the matte components by subtracting the error vector **e** from the actual luminance observation **y**. Note that the matte components obtained this way (Fig. 4, red solid line) almost coincide with the ground truth matte model, exhibiting a high degree of robustness.

4 Results and discussion

In this chapter, we present our experimental results and observations on synthesized and real datasets. All experiments were carried out on a Dell Optiplex 755 computer equipped with an Intel Core Duo E6550 CPU and 4 GB RAM, running Windows 7 Enterprise 64 bit. All algorithms were implemented in MATLAB R2014a 64 bit.

4.1 Normal map recovery

We first examine the angular error of normal maps recovered by different methods on both synthesized and real datasets. For synthesized datasets, we quantitatively inspect the difference between the ground truth normal map and the recovered normal maps. For real datasets without a ground truth map, on the other hand,
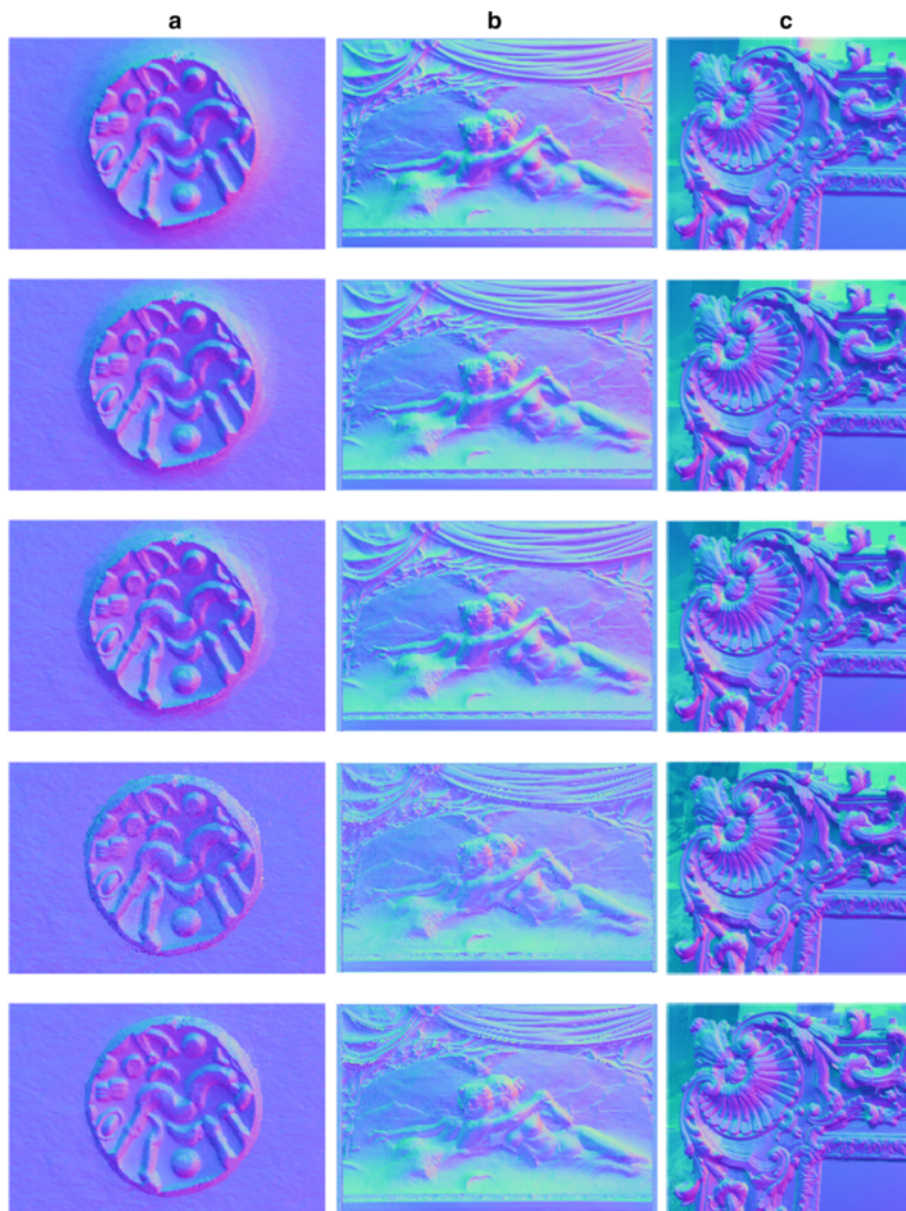

Fig. 12 Normal maps for three real-world dataset recovered using various methods. Columns **a–c** represent dataset *Gold*, *Elba* and *Frag*, respectively. Rows *1–5* show results using LS, IRL1, SBL, OMP, and LMS, respectively

the recovered normal maps are examined visually and qualitatively.

4.1.1 Synthesized datasets

Four 3D objects are used for our synthesized datasets in this study: *Sphere, Caesar, Buddha* and *Venus.* All 3D models are either created programmatically as geometrical primitives (*Sphere*) or downloaded from the AIM@SHAPE Shape Repository (*Caesar, Buddha*) [70] and the INRIA Gamma research database (*Venus*) [71]. For each object, 50 images are rendered under various lighting directions using raytracing software (POV-Ray 3.6) at a resolution of 200×200 (except for *Venus*, whose resolution is 150×250). Global illumination is enabled to ensure a highly photorealistic appearance. All scenes feature significant specularity and large areas of cast shadow. *Caesar, Buddha* and *Venus* are rendered with the specular highlight shading model provided by POV-Ray (a modified version of the Phong model) [72], and *Sphere* is rendered with a pure Phong model. A checkered plane is intentionally included in the rendered scene as background to (1) allow for the cast shadow to appear and (2) add further challenges to the algorithms since it introduces local fluctuation in luminance while the surface normals remain constant. Sample images for these datasets are shown in Fig. 5.
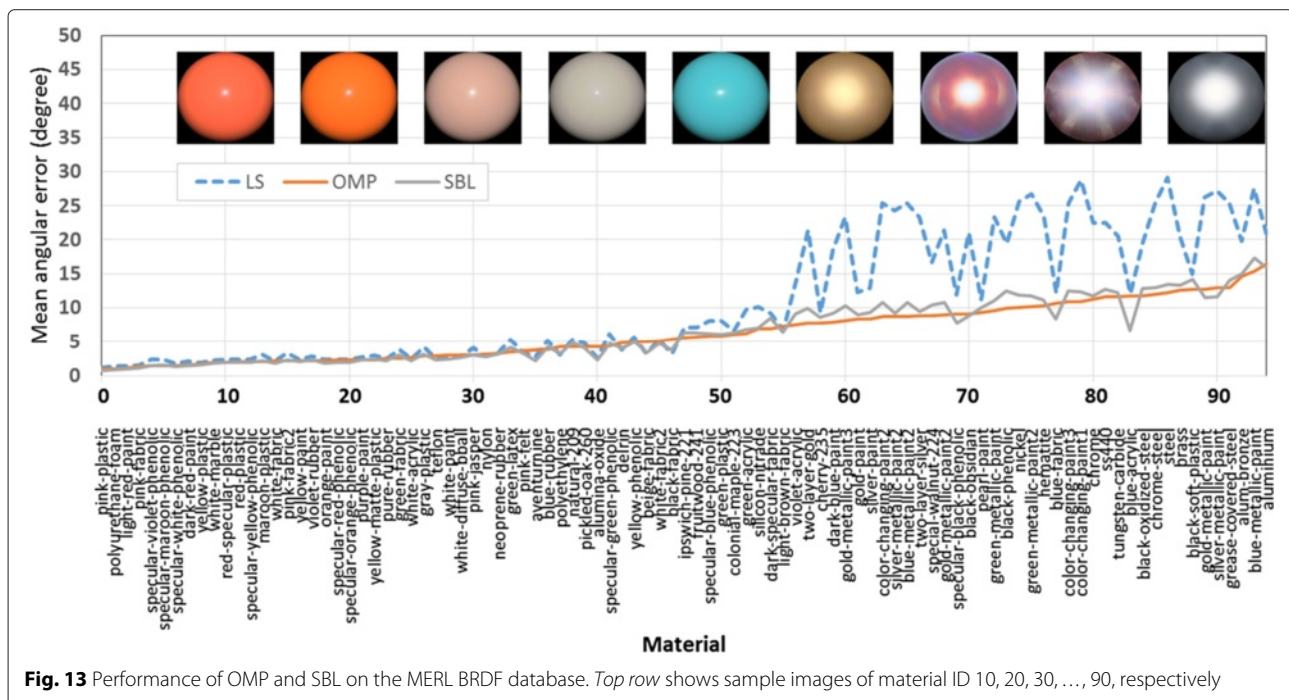
For each image set, the normal map is estimated using the OMP method [21–23] as proposed in this study. For comparison, we show the results for two other state-of-the-art sparse recovery methods—IRL1 and SBL [2].

Another two of our previously proposed outlier detection-based methods, LMS [10] and LMS mode finder [11], are also applied and compared. Then, the angular error between the normal map recovered using each method and the ground truth is quantitatively measured. Note that only results for *Caesar, Sphere* and *Buddha* are shown in this section. The fourth dataset *Venus* is reserved for later in Section 4.2.2 as a failure case.

We found these methods exhibit similar relative performance to each other on all the three datasets tested in this section—*Sphere, Caesar* and *Buddha*. In *Caesar*, the normal maps recovered using OMP (Fig. 6e) have a higher quality than those by IRL1 (Fig. 6g) and SBL (Fig. 6i) both qualitatively and quantitatively.

We observe that IRL1 and SBL, although much more robust than LS, still produce a considerable error at highly specular regions, most notably the cheek and the forehead. As a result, the faces on IRL1 and SBL normal maps appear to be more protruding than the ground truth. Also, some fine details on these two normal maps, such as the wrinkles on the forehead, are not well preserved. In addition, IRL1 and SBL fail to handle the regions right beside the neck which are heavily shadowed.

On the other hand, OMP shows a higher degree of robustness than previous sparse methods at specularity-affected regions (cheek, forehead and nose) as well as shadowed regions (areas around the neck on the checkered background), resulting in a normal map closer to the original. For example, the forehead appears flat on

Fig. 13 Performance of OMP and SBL on the MERL BRDF database. *Top row* shows sample images of material ID 10, 20, 30, ..., 90, respectively

OMP normal maps, closely resembling the ground truth. The wrinkles are almost perfectly recovered. However, OMP appears to be confused by the checkered pattern of background, producing a small angular error in these flat regions.

The LMS result (Fig. 6k) is better than IRL1 and SBL but worse than OMP.

The 1D version of LMS—finder—produces a poorer visual result (Fig. 6m) compared to the other robust methods, although it does give a statistically more reliable result than LS. We will exclude this method from future discussion but still show its result for reference.

The effect of specularity on normal map recovery can be further seen from the results for the *Sphere* dataset in Fig. 7. Again, IRL1 and SBL results are noisy in the specularity-affected areas, whereas OMP gives much cleaner results. LMS performs similarly to OMP. Interestingly, a pentagon-shaped pattern is visible on each error map because there are exactly five lights at each elevation angle.
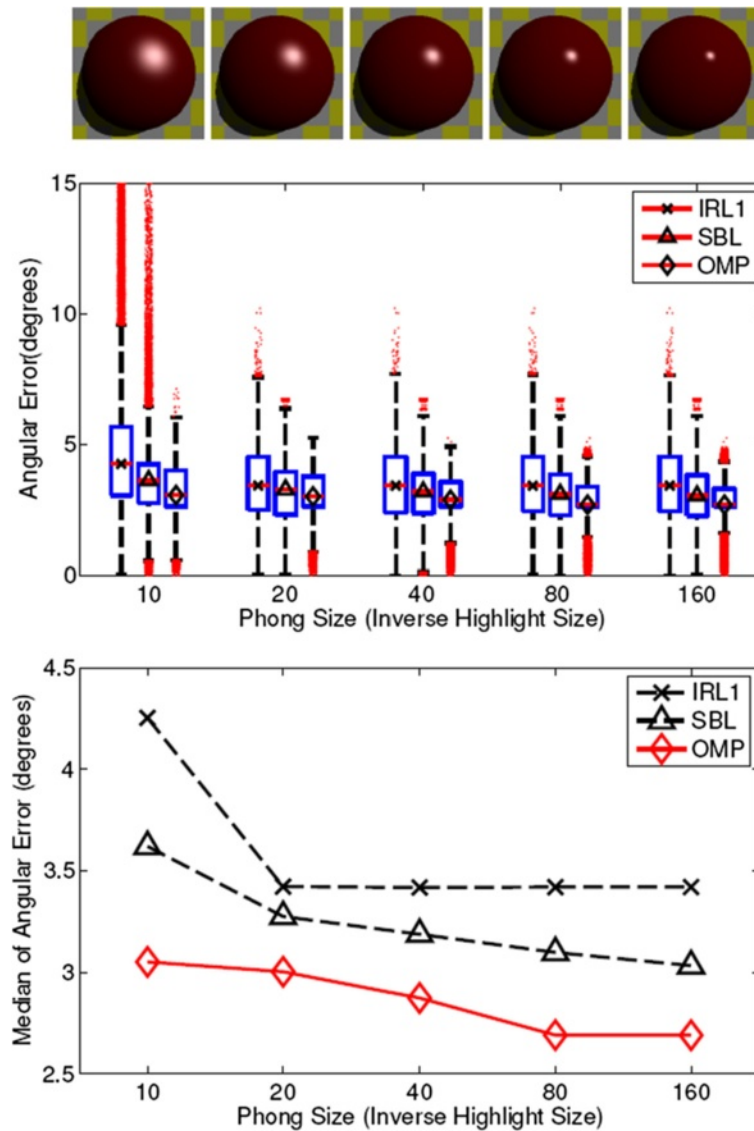

Fig. 14 Statistics for the angular error of normal maps recovered using sparse methods. *Upper row*: sample rendered images with *phong_size* 10–160, respectively, from left to right. *Middle figure*: boxplot of the angular error of normal maps under various degrees of specularity. *Upper and lower border of blue boxes* represent third (Q3) and first quartile (Q1), respectively. *Upper and lower whiskers* show 1.0 interquartile range (IQR) extended from Q3 to Q1, respectively. *Red bars in blue boxes* represent medians. At each specularity level, four sparse recovery methods are compared, symbolized by different markers on the median bar: IRL1 (*cross*), SBL (*triangle*) and OMP (*diamond*). Red dots are outliers distributed out of the error bar range. *Lower figure*: the medians of the angular error

For *Buddha* (Fig. 8), OMP again produces a better overall result than IRL1 and SBL. However, the relatively poorer performance of the greedy methods in shadowed concave regions now becomes a more significant problem due to the prevalence of concave regions such as creases on the clothes. The angular error distribution of the LMS result is similar to that of OMP, though with slightly greater overall error.

From the normal map recovery results obtained on the three datasets, we can see that OMP generally performs better than IRL1 and SBL on convex objects and are more resistant to specularities and cast shadows. The statistical result of the angular error of the normal maps recovered with different methods are listed in Table 1. OMP result has the lowest mean, median, 25 % and 75 % quantiles, as well as standard deviation for all three datasets mentioned above. The LMS result is better than IRL1 and SBL,
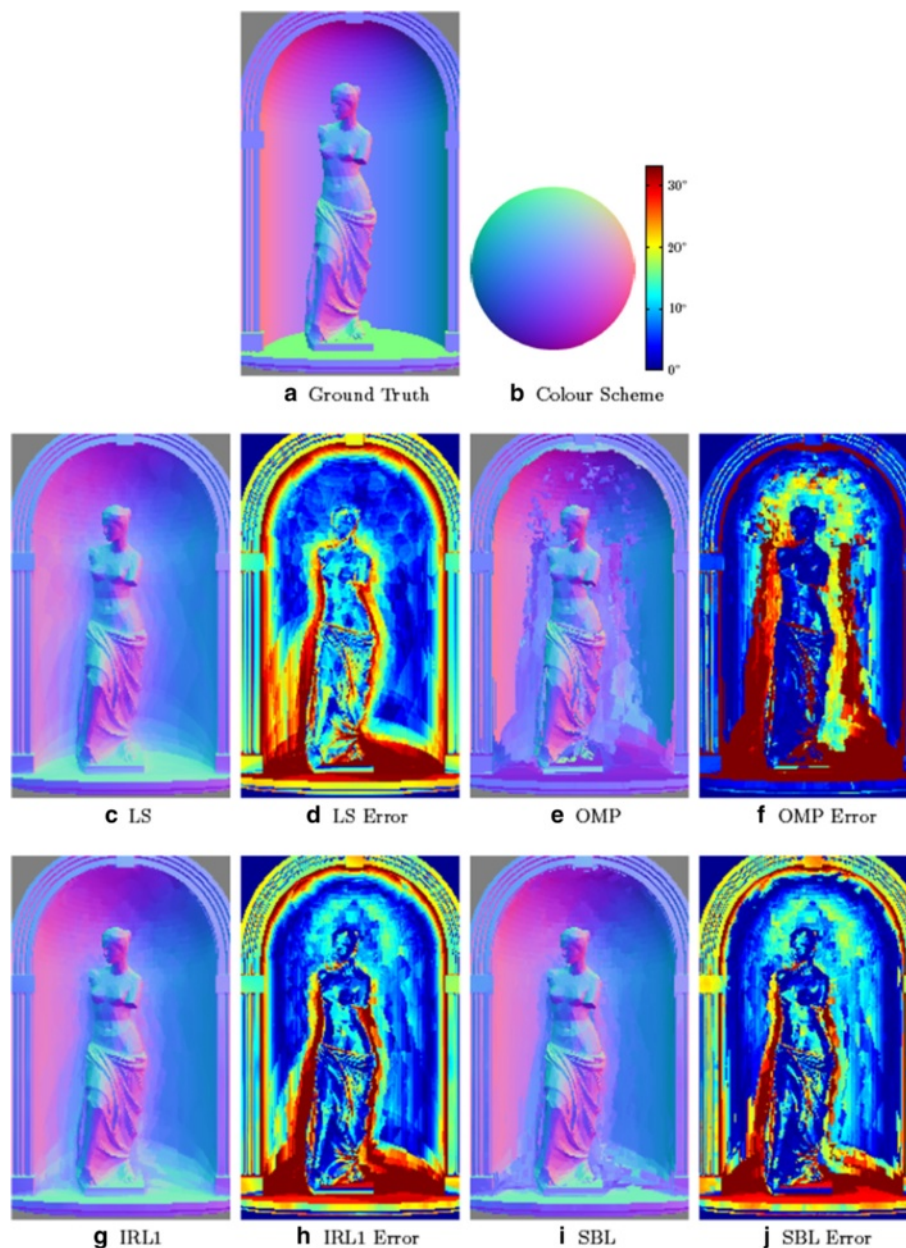

Fig. 15 Normal maps of the *Venus* dataset recovered using sparse methods. **a** Ground truth normal map. **b** Colour wheel and colour bar used for normal and angular error visualization, respectively. Angular error is measured in degrees. **c**, **e**, **g**, **i** Normal maps recovered using LS, OMP, IRL1 and SBL, respectively. **d**, **f**, **h**, **j** Angular error of normal maps recovered using the aforementioned methods

but worse than OMP. These results are also depicted in Fig. 9. Curiously, we notice that the estimation accuracy is generally lower on *Sphere* than *Caesar*, despite the simple geometry of the former. This may be jointly caused by the unique lighting model, surface colour and material that *Sphere* is rendered with. The exact explanation for this observation requires further investigation in the future.

As is witnessed on the *Buddha* dataset, OMP performs less optimally than IRL1 and SBL on small concave regions that are rarely illuminated. This problem also occurs for *Caesar* on the medial side of the eyes and under the

eyebrows. It is a lesser concern for objects that are generally convex such as *Caesar* and *Buddha* but may exert a strong negative influence on a scene that contains large concave areas. We will demonstrate the result for such a scene using *Venus* in Section 4.2.2.

4.1.2 Comparison via reconstructed surfaces

Using the normal maps recovered with various methods, we also reconstructed the 3D surface with the Frankot-Chellapa method [73] for direct comparison of the shape. Here, only the reconstruction result for *Caesar* is used
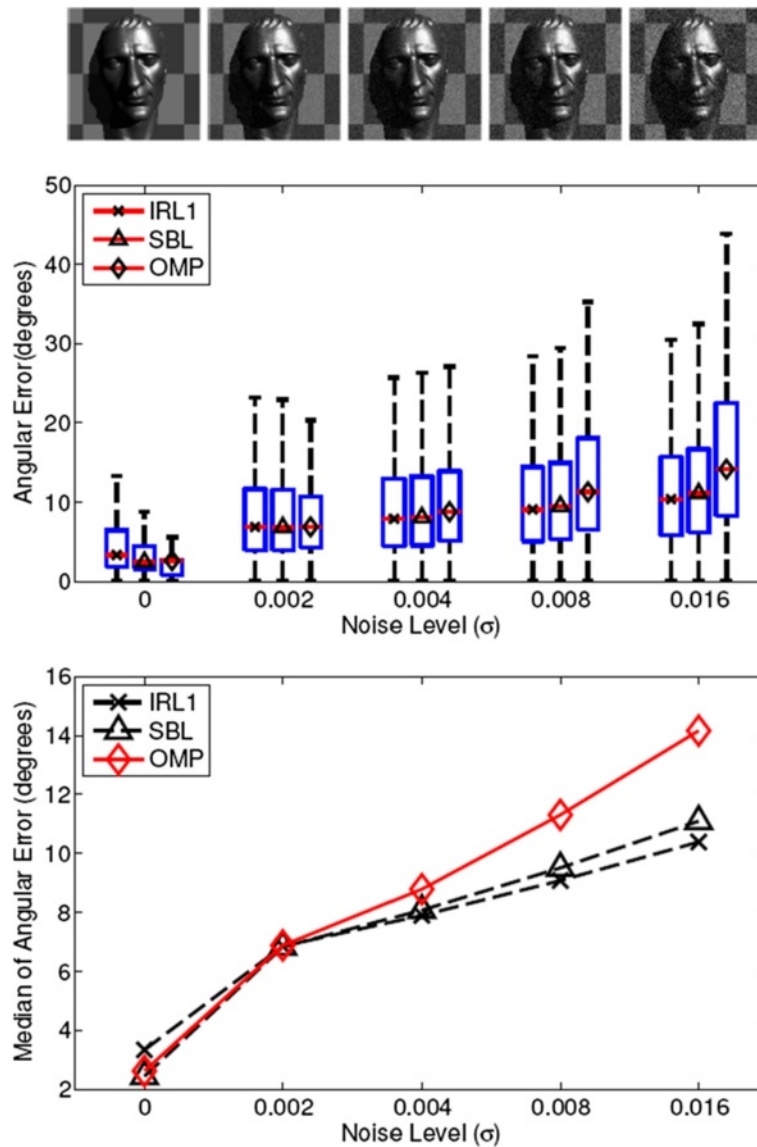

Fig. 16 Robustness of sparse recovery methods against Gaussian noise. *Upper tow*: sample images rendered with Gaussian noise of mean = 0.5 and STD = 0–0.016 , respectively, from left to right. *Middle figure*: boxplot of the angular error of normal maps under various degrees of Gaussian noise. Upper and lower border of *blue boxes* represent the third (Q3) and first (Q1) quartiles, respectively. *Upper and lower whiskers* show 1.0 IQR extended from Q3 and Q1, respectively. *Red bars in blue boxes* represent medians. At each noise level, three sparse recovery methods are compared, symbolized by different markers on the median bar: IRL1 (*cross*), SBL (*triangle*) and OMP (*diamond*). *Lower figure*: the medians of angular error

for demonstration. It is apparent from Fig. 10 that in the LS, IRL1, and SBL results, the overall shape of the face appear to be more protruding then it actually is, especially at the eyebrow ridge and the nose, whereas the OMP manages to preserve the shape accurately. Again, the LMS result appears to be less protruding than IRL1 and SBL results, although still not as accurate as OMP. We speculate that the exaggerated convexity originates from the inaccurately estimated normal vectors at highlight areas, such as the forehead and the nose. Since our greedy algorithm generally provides a better recovery in

those regions, they naturally yield a more accurate shape recovery.

4.1.3 Real datasets

Three datasets of real-world scenes are tested in this study: *Gold*, an ancient golden coin, *Elba*, an Italian high-relief sculpture, and *Frag*, a much-decorated golden frame (that surrounds a painting by Fragonard). Sample images of the three datasets are shown in Fig. 11.

The advantages and disadvantages of the methods we found using synthesized datasets are also observed in the
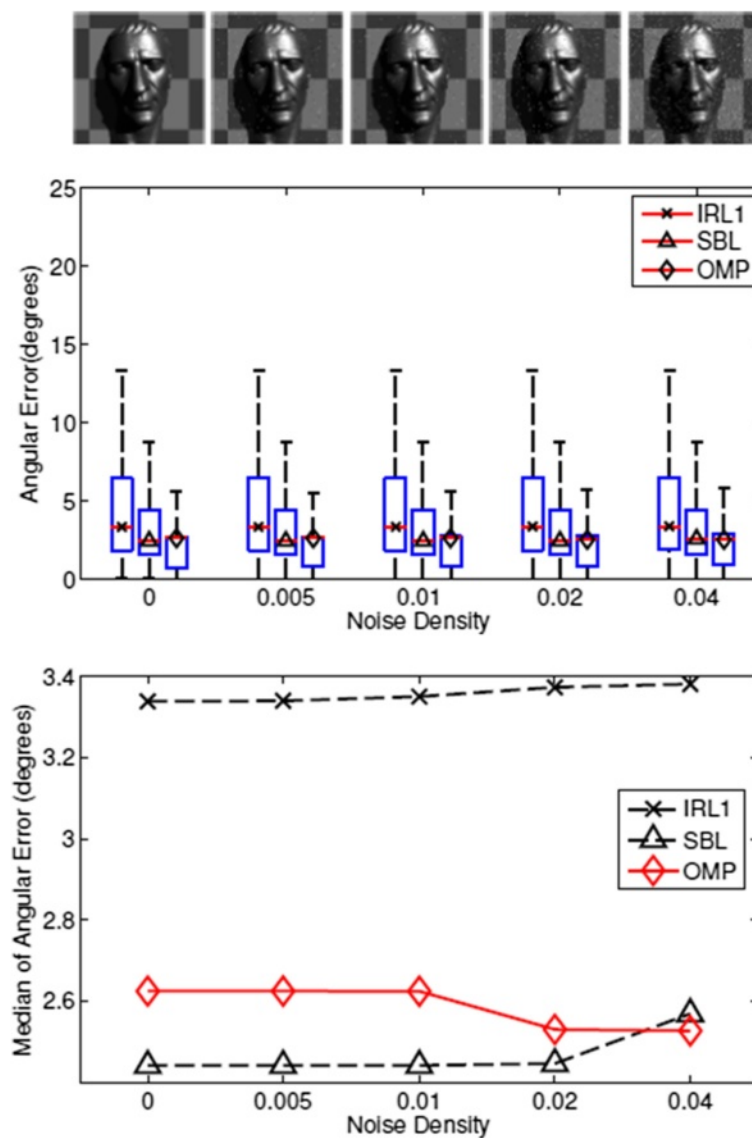

Fig. 17 Robustness of sparse recovery methods against salt and pepper noise. *Upper row*: sample images rendered with salt and pepper noise of density = 0–4 %, respectively, from left to right. *Middle figure*: boxplot of the angular error of normal maps under various degrees of salt and pepper noise. Upper and lower border of *blue boxes* represent third (Q3) and first (Q1) quartiles, respectively. *Upper and lower whiskers* show 1.0 IQR extended from Q3 and Q1, respectively. *Red bars in blue boxes* represent medians. At each noise level, four sparse recovery methods are compared, symbolized by different markers on the median bar: IRL1 (*cross*), SBL (*triangle*) and OMP (*diamond*). *Lower figure*: the medians of angular error

real datasets. Most images in dataset *Gold* have a large area of cast shadow. The influence of shadow can be clearly seen on the normal maps recovered by LS, IRL1 and SBL (Fig. 12a (1–3)) but is completely eliminated by OMP (Fig. 12a (4)). As for *Elba*, the scene contains a great number of small concave regions such as the pleats on the curtain. As expected, the greedy algorithm fail at these regions. Again, we notice that the LS, IRL1 and SBL results are more protruded than greedy results for both *Gold* and *Elba* (Fig. 12a (1–5), b (1–5)). Although there is not a ground truth normal map to support our speculation, it is reasonable to argue that the non-greedy algorithms exaggerate the convexity for *Elba*, as was the case for *Caesar*

(Fig. 10). The complex geometry of the object in our third dataset—*Frag*—accounts for the noisy estimates observed in concave regions in the greedy normal maps (Fig. 12C4 and C5). Note that the non-greedy results also show a large degree of inaccuracy in these regions (Fig. 12C1–C3), but in a less noticeable manner since these artefacts are usually smoothly blended into less-affected areas.

4.1.4 MERL database

We also tested the performance of OMP on 95 materials from the MERL BRDF database [74]. Each material is rendered on a sphere at a resolution of 200×200 on 50 images of various lighting directions. The performance
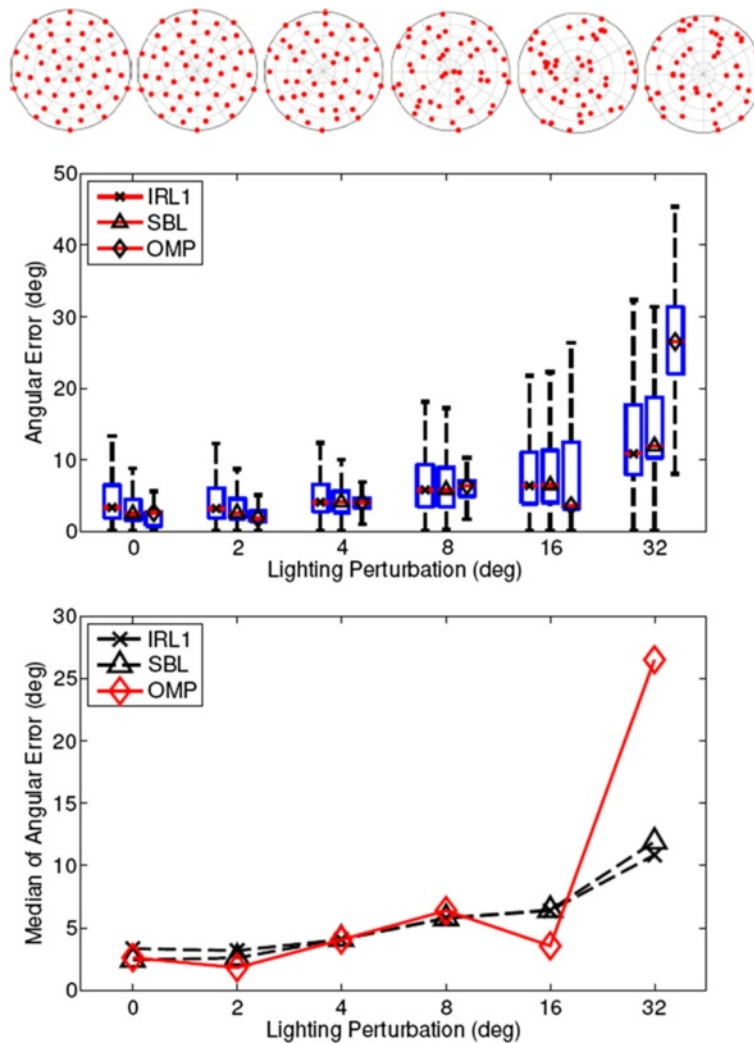

Fig. 18 Robustness of sparse recovery methods against light calibration error. *Top row*: leftmost plot shows actual light positions used for generating the dataset. Remaining plots show miscalibrated light positions with angular perturbations (2°–32°, from left to right) from the actual light positions at random directions. *Middle figure*: boxplot of the angular error of normal maps under various degrees of light calibration error. Upper and lower border of *blue boxes* represent third (Q3) and first (Q1) quartiles, respectively. *Upper and lower whiskers* show 1.0 IQR extended from Q3 and Q1, respectively. *Red bars in blue boxes* represent medians. At each angular perturbation level, four sparse recovery methods are compared, symbolized by different markers on the median bar: IRL1 (*cross*), SBL (*triangle*) and OMP (*diamond*). *Bottom figure*: the medians of angular error

pattern of OMP is very similar to SBL (Fig. 13): both methods are good at handling materials with an insignificant specular component (e.g. #10, *red-specular-plastic*). On the other hand, they both show decreased accuracy on shiny, strongly non-Lambertian metallic materials (e.g. #90, *silver-metallic-paint*) probably due to the violation of the sparsity assumption. Overall, the mean angular error of OMP over all 95 materials is 6.3174°, on par with SBL (6.5370°), and both methods significantly outperform the naive LS (10.8027°).

4.2 Robustness

To further understand of how well these methods behave in the presence of non-Lambertian effects, we tested their performance on *Sphere* with varying degrees of specularity and on *Venus,* where a large portion of the scene is concave, and as such, is heavily polluted by cast shadow. To find out the robustness of these methods against external error introduced by the experimental setup, we also tested the methods with additive image noise and light calibration error.

4.2.1 Specularity

We rendered five datasets of the same object *Sphere* with various sizes of highlight area (Fig. 14, top row) and tested how the size of the specular region affects the performance of our sparse regression methods. The size of the highlight is controlled by the *phong_size* parameter in POV-Ray [72]. We found that although the accuracy of all three methods compared (IRL1, SBL, OMP) decreases as the specular size increases, the greedy algorithm is less affected (Fig. 14, middle and bottom figures).

4.2.2 Shadow and concavity: a failure case

In Section 4.1.1, we have already noticed the possibility that the performance of our greedy algorithm may be negatively affected at shadowed concave regions. Here, we use the *Venus* dataset to further demonstrate this observation. In *Venus* (Fig. 5, bottom row), the convex foreground (the Venus statue) and the concave background (the dome) are well separated, allowing us to clearly inspect the performance of algorithms on different regions.

The result is shown in Fig. 15. As speculated, OMP shows robustness in shiny, convex regions such as the outer rim of the dome, and on the statue itself, but fails on the heavily shadowed background. The other three methods (LS, IRL1, SBL), on the contrary, suffer from noticeable angular error in convex areas. However, they are less severely affected by shadow and concavity on the background than greedy methods. Overall, the normal map recovered with the greedy approach are less smooth for *Venus* due to the inaccurate estimation of normal vectors in the concave regions.

4.2.3 Image noise

We tested three sparse algorithms (IRL1, SBL and OMP) against Gaussian noise as well as salt and pepper noise. For Gaussian noise (Fig. 16), the accuracy of all three methods drastically decreases as the noise level increases, although OMP appears to be slightly more adversely affected. On the other hand, all sparse methods are quite insensitive to salt and pepper noise (Fig. 17).

4.2.4 Lighting

There might be cases when the lighting directions are not properly calibrated. That is, the assumed lighting directions deviate from their actual values. In this test, we introduce for every assumed lighting vector a fixed angular perturbation, ranging from 2° to 32°, at a random direction, while keeping the actual arrangement of lights unchanged.

We tested the performance of the sparse methods under various degrees of light calibration error on the *Caesar* dataset. The actual arrangement of lights is displayed in Fig. 18 (leftmost plot on the top row). As an increasingly greater random perturbation is added to the assumed lighting directions, the angular error gradually increases for all sparse methods. Note that OMP appears to be susceptible to the random calibration error the most, especially when the perturbation reaches 32°.

Also note that in Fig. 18 (bottom), the median of the angular error produced by OMP slightly decreases at 16° compared to previous conditions. We believe that this is a fluctuation caused by the particular arrangement of lights at this condition. Despite this decrease in the median of error, the widths of the error distributions steadily increase at 16° for all three methods, as can be clearly seen from Fig. 18 (middle).

It was reported that the number of lights has a large impact on the accuracy of sparse photometric stereo recovery [2]. We found that our OMP-based method also shows a similar but somewhat greater dependency on the
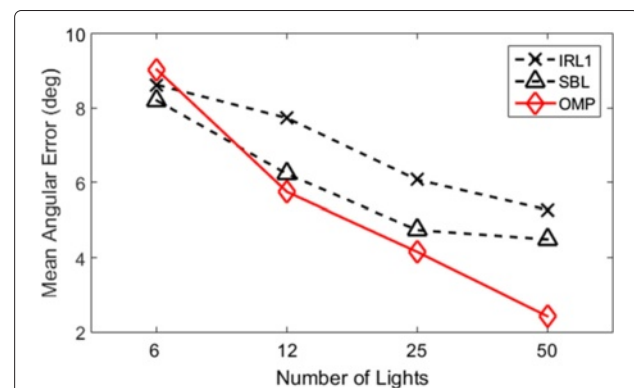

Fig. 19 Mean angular error of recovered normal maps under different numbers of lights

number of lights (Fig. 19). This observation indicates that the OMP works best when a large number of lights are present.

We also investigated the performance of sparse methods under a highly biased lighting distribution. We used 25 lights; 23 of them is located on the left or upper-left hemisphere and the other two on the right (Fig. 20). Under such a biased lighting, OMP still has the best mean angular error (8.2140°) compared with IRL1 (9.1173°), SBL (8.6561°) and LS (12.0726°).

4.3 Efficiency

The actual per-pixel processing time for the MATLAB implementation of the algorithms tested in this study is reported in Fig. 21. The maximum number of iterations for IRL1 and SBL are set to 100 although iteration will be terminated as soon as another stopping criterion is met; OMP always terminates after exactly 28 iterations for our datasets of 50 images; for LMS, the number of iterations is fixed at 1500.

In OMP, the operation with the highest asymptotic complexity is the inversion of a $k \times k$ matrix (where k is the number of selected columns) in Eq. 14. With a naive Gauss-Jordan elimination method, the inversion takes $O(k^3)$, which is asymptotically $O(n^3)$ since $k \le n/2 + 3$. Since the above operation is repeated $n/2 + 3$ times, the overall time complexity of our OMP algorithm is $O(n^4)$. In our current implementation, the running time of OMP (4.823 ms/pixel) is comparable to IRL1 (3.338 ms/pixel). LMS is the slowest (57.48 ms/pixel), though it can be made faster with fewer iterations at the expense of accuracy.
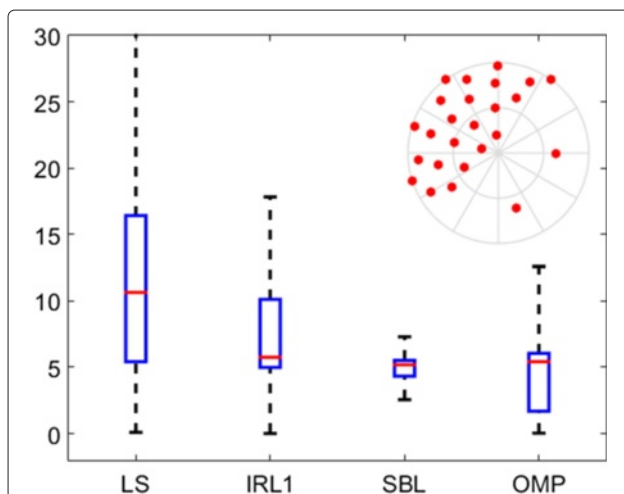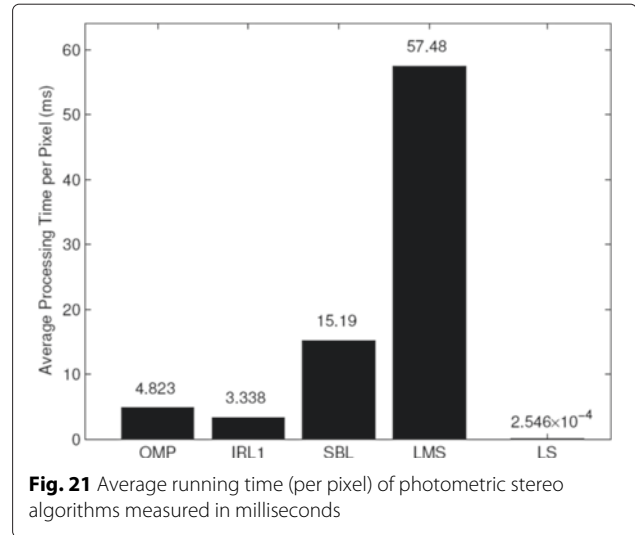

Fig. 21 Average running time (per pixel) of photometric stereo algorithms measured in milliseconds

4.4 Summary

Based on the experimental results above, we have come to the conclusion that our greedy algorithm overall has a higher accuracy than L1 minimization and SBL with a comparable efficiency, though OMP may be less robust in poorly illuminated regions. LMS is close to the greedy sparse algorithm in accuracy, despite its low efficiency. The algorithms tested in this chapter are summarized and compared in Table 2.

5 Conclusions

In this study, the classical PST is reformulated in terms of the canonical form of sparse recovery, and a greedy algorithm—OMP—is applied to solve the problem. Our formulation is different from previous ones [1, 2] in that the former incorporates normal vector components and non-Lambertian errors in one combined vector, allowing for the straightforward application of OMP. In order for OMP to obtain normal estimations, the normal vector components have to be selected before the iteration stops. Although it is not theoretically guaranteed, we observed that the normal components are always selected within the first few iterations in the datasets we tested, unless some components are indeed zero or very close to zero. We also speculate that the dictionary matrix in our formulation is near-orthonormal and satisfies the conditions required by OMP to achieve exact recovery.

We found that, in general, our greedy method OMP outperforms other state-of-the-art sparse solvers such as IRL1 and SBL [2] with comparable efficiency. In particular, OMP provides a more numerically accurate estimation of normal vectors in the presence of common non-Lambertian effects such as highlights and cast shadows, although it may occasionally fail at concave areas that are poorly illuminated. In addition, all sparse methods

Fig. 20 Recovery accuracy under highly biased lighting distribution. Polar plot shows the arrangement of lights. Upper and lower border of *blue boxes* represent third (Q3) and first (Q1) quartiles, respectively. *Upper and lower whiskers* show 1.0 IQR extended from Q3 and Q1, respectively. *Red bars in blue boxes* represent medians

Table 2 Qualitative comparison of photometric stereo algorithms

	Robustness				Smoothness	Efficiency
	Overall	Highlight	Shadow	Concavity		
LS	Very low	Very low	Very low	Very low	Very high	Very high
LMS	Very high	Very high	Very high	–	Low	Very low
LMS mode	Low	Low	Low	–	Medium	Very high
IRL1	High	High	High	Medium	High	High
SBL	High	High	High	Medium	High	Low
OMP	Very high	Very high	Very high	Low	Medium	High

The performance is evaluated on a five-level scale: 'Very low', 'low', 'medium', 'high' and 'very high'. Fields that are not available are indicated by a '–' sign

tested are reasonably robust against additive image noise and lighting calibration error.

Another two outlier-removal based methods—LMS and LMS mode finder—are also tested in this study for comparison. LMS results are overall statistically more accurate than IRL1 and SBL but less so than OMP. LMS mode finder, the 1D simplification of LMS, shows some robustness against non-Lambertian errors, especially highlights but performs poorly against cast shadows.

This study opens up many possible directions for future research. First, a great number of sparse recovery algorithms have already been proposed in the past few decades, each designed for a specific formulation. Even within the domain of greedy algorithms, there are many other potential candidates aside from OMP that may be directly applied to the PST problem. It would be interesting to explore this large repertoire of sparse formulations and recovery algorithms to find an optimal method.

It has been shown that sparse methods such as IRL1 and SBL can be used to estimate the lighting directions in the context of uncalibrated PST [19]. It is highly possible that greedy algorithms such as OMP can also be extended to be applied for such a purpose. Future studies may reveal more applications of greedy algorithms in different aspects of the PST framework.

References
1. L Wu, A Ganesh, B Shi, Y Matsushita, Y Wang, Y Ma, in *Proceedings of Asian Conference of Computer Vision*. Robust photometric stereo via low-rank matrix completion and recovery (Springer, Berlin Heidelberg, 2010), pp. 703–717
2. S Ikehata, D Wipf, Y Matsushita, K Aizawa, in *IEEE Conference on Computer Vision and Pattern Recognition*. Robust photometric stereo using sparse regression (IEEE Computer Society, Washington DC, USA, 2012), pp. 318–325
3. G Fyffe, X Yu, P Debevec, in *Int. Conf. on Computational Photog*. Single-shot photometric stereo by spectral multiplexing (IEEE Computer Society, Washington DC, USA, 2011), pp. 1–6
4. G Fyffe, P Debevec, in *Int. Conf. on Computational Photog*. Single-shot reflectance measurement from polarized color gradient illumination (IEEE Computer Society, Washington DC, USA, 2015)
5. A Hertzmann, SM Seitz, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Shape and materials by example: a photometric stereo approach, vol. 1 (IEEE Computer Society, Washington DC, USA, 2003), pp. 533–540
6. N Alldrin, T Zickler, D Kriegman, in *IEEE Conference on Computer Vision and Pattern Recognition*. Photometric stereo with non-parametric and spatially-varying reflectance (IEEE, Anchorage AK, 2008), pp. 1–8
7. J Ackermann, M Ritz, A Stork, M Goesele, in *Trends and Topics in Computer Vision,* Lecture Notes in Computer Science, ed. by KN Kutulakos. Removing the example from example-based photometric stereo (Springer, Berlin Heidelberg, 2012), pp. 197–210
8. F Verbiest, L Van Gool, in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Photometric stereo with coherent outlier handling and confidence estimation (IEEE Computer Society, Washington DC, USA, 2008), pp. 1–8
9. C Hernandez, G Vogiatzis, R Cipolla, Multiview photometric stereo. IEEE Trans. Pattern Anal. Mach. Intell. **30**(3), 548–554 (2008)
10. MS Drew, Y Hel-Or, T Malzbender, N Hajari, Robust estimation of surface properties and interpolation of shadow/specularity components. Image Vis. Comput. **30**(4–5), 317–331 (2012)
11. M Zhang, MS Drew, in *CPCV2012: European Conference on Computer Vision Workshop on Color and Photometry in Computer*. Vision Lecture Notes in Computer Science. Robust luminance and chromaticity for matte regression in polynomial texture mapping, vol. 7584/2012 (Springer, Berlin Heidelberg, 2012), pp. 360–369
12. MS Drew, N Hajari, Y Hel-Or, T Malzbender, in *Proceedings of the British Machine Vision Conference*. Specularity and shadow interpolation via robust polynomial texture maps, (2009), pp. 114–111411
13. DL Donoho, Compressed sensing. IEEE Trans. Inf. Theory. **52**(4), 1289–1306 (2006)
14. BK Natarajan, Sparse approximation solutions to linear systems. SIAM J. Comput. **24**(2), 227–234 (1995)
15. RJ Woodham, Photometric method for determining surface orientation from multiple images. Opt. Eng. **19**(1), 139–144 (1980)
16. S Barsky, M Petrou, The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. IEEE Trans. Pattern Anal. Mach Intell. **25**(10), 1239–1252 (2003)
17. H Rushmeier, G Taubin, A Guéziec, in *Proceedings of Eurographics Workshop on Rendering*. Applying shape from lighting variation to bump map capture, (1997), pp. 35–44
18. P Favaro, T Papadhimitri, A closed-form solution to uncalibrated photometric stereo via diffuse maxima. IEEE Conf. Computer Vision and Pattern Recognit. **157**(10), 821–828 (2012)
19. V Argyriou, S Zafeiriou, B Villarini, M Petrou, A sparse representation method for determining the optimal illumination directions in photometric stereo. Signal Process. **93**(11), 3027–3038 (2013)

20. J Wright, AY Yang, A Ganesh, SS Sastry, Y Ma, Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 210–227 (2009)

21. YC Pati, R Rezaiifar, PS Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. Proceedings of 27th Asilomar Conference on Signals, Systems and Computers. **1**, 40–44 (1993)

22. JA Tropp, AC Gilbert, Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. Inf. Theory. **53**(12), 4655–4666 (2007)

23. TT Cai, L Wang, Orthogonal matching pursuit for sparse signal recovery with noise. IEEE Trans. Inf. Theory. **57**(7), 4680–4688 (2011)

24. EN Coleman Jr., R Jain, Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. Comput. Graph. Image Process. **18**, 309–328 (1982)

25. F Solomon, K Ikeuchi, Extracting the shape and roughness of specular lobe objects using four light photometric stereo. IEEE Trans. Pattern Anal. Mach. Intell. **18**, 449–454 (1996)

26. A Yuille, D Snow, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Shape and albedo from multiple images using integrability (IEEE Computer Society, Washington DC,USA, 1997), pp. 158–164

27. G Willems, F Verbiest, W Moreau, H Hameeuw, K Van Lerberghe, L Van Gool, in *Short and Project Papers Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*. Easy and cost-effective cuneiform digitizing (Eurographics Association, Geneve, Switzerland, 2005), pp. 73–80

28. J Sun, M Smith, L Smith, S Midha, J Bamber, Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities. Image Vis. Comput. **25**, 1050–1057 (2007)

29. S Mallick, T Zickler, D Kriegman, P Belhumeur, in *IEEE Comp. Soc. Conf. on Comp. Vis. and Patt. Rec. 2005*. Beyond lambert: reconstructing specular surfaces using color, vol. 2, (2005), pp. 619–6262

30. M Holroyd, J Lawrence, G Humphreys, T Zickler, A photometric approach for estimating normals and tangents. ACM Trans. Graph. (Proceedings of SIGGRAPH Asia 2008). **27**(5), 32–39 (2008)

31. C Julià, F Lumbreras, AD Sappa, A factorization-based approach to photometric stereo. Int. J. Imaging Syst. Technol. **21**(1), 115–119 (2011)

32. K-L Tang, C-K Tang, T-T Wong, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Dense photometric stereo using tensorial belief propagation, vol. 1, (2005), pp. 132–139

33. M Chandraker, S Agarwal, D Kriegman, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. ShadowCuts: photometric stereo with shadows, (2007), pp. 1–8

34. Y Mukaigawa, Y Ishii, T Shakunaga, Analysis of photometric factors based on photometric linearization. J. Opt. Soc. Am. A, Optics, image science, and vision. **24**(10), 3326–3334 (2007)

35. D Miyazaki, K Hara, K Ikeuchi, Median photometric stereo as applied to the Segonko Tumulus and museum objects. Int. J. Comput. Vision. **86**(2–3), 229–242 (2010)

36. HD Tagare, RJP de Figueiredo, A theory of photometric stereo for a class of diffuse non-Lambertian surfaces. IEEE Trans. Pattern Anal. Mach. Intell. **13**(2), 133–152 (1991)

37. SK Nayar, K Ikeuchi, T Kanade, Determining shape and reflectance of hybrid surfaces by photometric sampling. IEEE Trans. Robot. Autom. **6**(4), 418–431 (1990)

38. AS Georghiades, in *9th IEEE International Conference on Computer Vision*. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo, (2003), pp. 816–8232. doi:10.1109/ICCV.2003.1238432

39. T Malzbender, D Gelb, H Wolters, in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '01. Polynomial texture maps (ACM, New York, USA, 2001), pp. 519–528

40. R Basri, D Jacobs, I Kemelmacher, Photometric stereo with general, unknown lighting. Int. J. Comput. Vision. **72**, 239–257 (2007)

41. WM Silver, *Determining shape and reflectance using multiple images*. (Massachusetts Institute of Technology, Cambridge MA, 1980)

42. SJ Koppal, SG Narasimhan, in *IEEE Conference on Computer Vision and Pattern Recognition*. Clustering Appearance for Scene Analysis, vol. 2, (2006), pp. 1323–1330

43. D Goldman, B Curless, A Hertzmann, S Seitz, in *10th IEEE International Conference on Computer Vision*. Shape and spatially-varying BRDFs from photometric stereo, vol. 1, (2005), pp. 341–348

44. Q Yang, N Ahuja, Surface reflectance and normal estimation from photometric stereo. Comput. Vision and Image Underst. **116**(7), 793–802 (2012)

45. S Kherada, P Pandey, A Namboodiri, in *IEEE Workshop on Applications of Computer Vision*. Improving realism of 3D texture using component based modeling (IEEE Computer Society, Washington DC, USA, 2012), pp. 41–47

46. B Shi, T Ping, Y Matsushita, K Ikeuchi, Bi-polynomial modeling of low-frequency reflectances. IEEE Trans. Pattern Anal. Mach. Intell. **36**(6), 1–1 (2013)

47. Z Lin, M Chen, Y Ma, *The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices*. (Technical report, UIUC (UILU-ENG-09-2215), 2009)

48. E Candès, J Romberg, l1-magic: recovery of sparse signals via convex programming. URL: http://statweb.stanford.edu/~candes/l1magic/downloads/l1magic.pdf 4(2005)

49. EJ Candès, MB Wakin, SP Boyd, Enhancing sparsity by reweighted L1 minimization. J. Fourier Anal. Appl. **14**(5–6), 877–905 (2008)

50. DP Wipf, BD Rao, Sparse Bayesian learning for basis selection. IEEE Trans. Signal Process. **52**(8), 2153–2164 (2004)

51. A Adler, M Elad, Y Hel-Or, E Rivlin, Sparse coding with anomaly detection. J. of Signal Proc. Systems. **79**, 179–188 (2015)

52. SP Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, Cambridge, UK, 2004), p. 744

53. MAT Figueiredo, RD Nowak, SJ Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. IEEE J. Selected Topics in Signal Process. **1**(4), 586–597 (2007)

54. SG Mallat, Z Zhang, Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process. **41**(12), 3397–3415 (1993)

55. D Needell, R Vershynin, Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. Found. Comput. Math. **9**(3), 317–334 (2009)

56. D Needell, R Vershynin, Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. IEEE J. Selected Topics in Signal Process. **4**(2), 310–316 (2010)

57. DL Donoho, Y Tsaig, I Drori, J-L Starck, Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. IEEE Trans. Inf. Theory. **58**(2), 1094–1121 (2012)

58. D Needell, JA Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples. Commun. ACM. **53**(12), 93–100 (2010)

59. A Divekar, O Ersoy, *Probabilistic matching pursuit for compressive sensing*. (Technical report, Purdue University, West Lafayette, IN, USA, 2010)

60. S Chatterjee, D Sundman, M Skoglund, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Look ahead orthogonal matching pursuit, (2011), pp. 4024–4027

61. P Jain, A Tewari, IS Dhillon, in *Proceedings of Neural Information Processing Systems*. Orthogonal matching pursuit with replacement, (2011), pp. 1215–1223

62. NB Karahanoglu, H Erdogan, A* orthogonal matching pursuit: best-first search for compressed sensing signal recovery. Digital Signal Process. **22**(4), 555–568 (2012)

63. T Blumensath, ME Davies, Iterative hard thresholding for compressed sensing. Appl. Comput. Harmonic Anal. **27**(3), 265–274 (2009)

64. W Dai, O Milenkovic, Subspace pursuit for compressive sensing signal reconstruction. IEEE Trans. Inf. Theory. **55**(5), 2230–2249 (2009)

65. DL Donoho, A Maleki, A Montanari, Message-passing algorithms for compressed sensing. Proc. Natl. Acad. Sci. **106**(45), 18914–18919 (2009)

66. A Maleki, DL Donoho, Optimally tuned iterative reconstruction algorithms for compressed sensing. IEEE J. Sel. Topics in Signal Process. **4**(2), 330–341 (2010)

67. V Cevher, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. An ALPS view of sparse recovery, (2011), pp. 5808–5811

68. EJ Candès, T Tao, Decoding by linear programming. IEEE Trans. Inf. Theory. **51**(12), 4203–4215 (2005)

69. EJ Candès, JK Romberg, T Tao, Stable signal recovery from incomplete and inaccurate measurements. Commun.Pure and Appl. Math. **59**(8), 1207–1223 (2006)

70. AIM@SHAPE shape repository. http://visionair.ge.imati.cnr.it/ontologies/shapes/
71. 3D meshes research databse by INRIA Gamma Group. http://www-roc.inria.fr/gamma/gamma/download/download.php
72. POV-Ray 3.6.0 documentation online. http://www.povray.org/documentation/view/3.6.0/347/
73. RT Frankot, R Chellapa, A method for enforcing integrability in shape from shading algorithms. IEEE Trans. Pattern Anal. Mach. Intell. **10**(4), 439–451 (1988)
74. W Matusik, H Pfister, M Brand, L McMillan, A data-driven reflectance model. ACM Trans. Graph. **22**(3), 759–769 (2003)